HL7 RCRIM WG

HL7 Version 3 Standard: Regulated Studies; Clinical Research Filtered Query (CRFQ) Service Functional Model (SFM)

Draft Standard for Trial Use, Release 1

May 2008

Publication of this draft standard for trial use and comment has been approved by Health Level Seven, Inc. (HL7). Distribution of this draft standard for comment shall not continue beyond 24 months from the date of publication. It is expected that following this 24 month period, this draft standard, revised as necessary, will be submitted to a normative ballot in preparation for approval by ANSI as an American National Standard. This draft standard is not an accredited American National Standard. Suggestions for revision should be submitted at http://www.hl7.org/dstucomments/index.cfm.

Service Functional Model Specification (SFM)

Clinical Research Filtered Query (CRFQ)

Version 1.3.03

April 30, 2008

Project Lead	Charlie Mead, Booz Allen Hamilton (BAH)/National Cancer Institute (NCI)
Editor	Charlie Mead, BAH/NCI
Authors	John Koisch, BAH Charlie Mead, BAH/NCI Marti Velezis, BAH
Key Contributors	Isabelle Dezegher, Novartis Joyce Niland, City of Hope Mitra Rocca, Novartis Denise Wold, GSK
Contributors	Lisa Chattergee, Digital Infuzion Isabelle Dezegher, Novartis Bill Friggle, Sanofi-Aventis Terry Hardin, IBM Barbara McKinnon Joyce Niland, City of Hope Mitra Rocca, Novartis

Preface

Note to Readers

This document is the Service Functional Model (SFM) for the Clinical Research Filtered Query (CRFQ) service. The general content of this document is specified under the Service Specification Framework (SSF) process under the auspices of the Healthcare Services Specification Project (HSSP). Readers of this document should keep in mind that the SFM provides a Service *interface specification*, *not* a specification of a Service *implementation*. (A more detailed contextual discussion of this point is given in the Overview section of this document.) The distinction between a service interface specification and a service implementation is a critical point of distinction in context of Service Oriented Architecture (SOA) methodologies and implementations. In particular, a single service interface specification can – and often does – have more than one associated service implementation, i.e. there are different, semantically consistent ways of implementing all or part of the functionality specified in the SFM and therefore to support the behavior described in this SFM/specification.

Changes from Previous Release of the CRFQ service

This is the first SFM for the CRFQ. Hence, there are no changes from a previous release.

Acknowledgements

The BRIDG Project

Health Level 7 (HL7)

Health Service Specification Project (HSSP)

Object Management Group (OMG)

The ASPIRE Project

Guide to Readers

A brief description of each of the Sections of this document is shown in the table below:

Section	Description
1	Describes services in healthcare, HSSP, and the HSSP process
2	Describes the real-world representation of the service, from specification through implementation, with a focus on market-place relevance
3	Business Cases and Storyboards
4	Dependencies and Assumptions
5	Functional Model, including elaborated operations
6	Functional groupings of operations to respond to business focuses
7	Business Scenarios depicted in behavioral diagrams
8	The services framework functional model
9	The relationship between information and interface functionality
10	Recommendations from the HL7 community to the responders to the RFP

Readers may want to target to review of this document based on their perspective of the CRFQ service. The following table provides a high-level guide for readers based on their association with:

- the Regulated Clinical Research Information Management Technical Committee or other organizations, groups, or committees consisting primarily of Subject Matter Experts (SMEs);
- architecture- or design-level groups such as HL7SOA or HSSP; and/or
- responders to the OMG RFP that will be issued once the CRFQ becomes an ANSI Draft Standard for Trial Use (DSTU) through the HL7-supervised ANSI balloting process, as well as other institutions or organizations that may be interested in building initial implementation of the CRFQ service independent of the OMG RFP process.

Audience	**Sections (in order of Priority)**
RCRIM TC / SME's	2, 3, 10
SOA4HL7 TC / HSSP / architects / designers	6, 5, 10, 7, 4
RFP Submitters / other implementers	2, 10, 3, 7, 5

Table of Contents

108

1 Overview

1.1 Introduction and Scope

The Service Specification Framework (SSF) is the methodology developed by the HSSP as the prescribed process to be followed when defining specific HSSP-compliant service specifications. The methodology sets out an overall process including specifying a template for the content of the Service Functional Model (SFM). Section 2 of the SFM describes the business context for the service that is the focus of each SFM document (in this SFM, the CRFQ service). The SSF dictates, however, that a preamble to Section 2 describes the larger 'HSSP context' of the target service and its specification, i.e. 'a view of the service from a methodological perspective.'

1.1.1 HL7-OMG Healthcare Services Specification Project (HSSP)

The Healthcare Services Specification Project (HSSP) [http://hssp.wikispaces.com] is a joint endeavor between Health Level Seven (HL7) [http://www.hl7.org] and the Object Management Group (OMG) [http://www.omg.org]. The HSSP was chartered at the January 2005 HL7 meeting under the Electronic Health Records Technical Committee (EHR TC), and the project was subsequently approved by the Board of Directors of both HL7 and OMG.

The HSSP has several objectives including:

- Stimulating the adoption and use of standardized "plug-and-play" services by healthcare software product vendors;

- Facilitation of the development of a set of implementable interface standards supporting agreed-upon services specifications which will collectively form the basis for provider purchasing and procurement decisions; and

- Development of a complementary service specification framework around which clinical-care-, clinical-research-, and life-science-focused business and infrastructure services can be defined in such a way as to leverage (rather than conflict or compete with) existing HL7 work products and processes, thereby leveraging content and process ;lessons learned' from elsewhere within the organization's various TCs and Special Interest Groups (SIGs).

Within the SSF processes, HL7 has primary responsibility for:

- identifying and prioritizing services as candidates for standardization;
- specifying the functional requirements and conformance criteria for these services in the form of Service Functional Model (SFM) specifications such as this document; and
- adopting these SFMs as balloted HL7/ANSI balloted standards.

These activities are coordinated by the HL7 Services Oriented Architecture Technical Committee in collaboration with other HL7 SIGs and TCs, the involved groups depending on the focus of the particular service specification being developed. Historical participants have included the Vocabulary TC and Clinical Decision Support SIG. The CRFQ is a sponsored project of the Regulated Clinical Research Information Management (RCRIM) TC.

Once an SFM developed within HL7 becomes a balloted ANSI Draft Standard for Trial Use (note that all SFMs produced by HL7 are balloted as Draft Standards for Trial Use (DSTUs) to enable fluid evolution of the service specification post-implementation and testing), the Object Management Group (OMG) issues a "Request for Proposal" (RFP) -- for implementations of the service based on the SFM. This process is the basis for OMG's industry-driven standards-development process, a process that allows vendors and other submitters to propose solutions that satisfy the mandatory and optional requirements expressed in the RFP (and derived from the SFM), while leaving design flexibility to the submitters and implementation flexibility to the users of the standard. The result of this collaboration is an RFP Submission, which will be referred to in the HSSP process as a Service Technical Model (STM). HL7 members' concerns regarding the content and functionality are explicitly included in during the RFP development process as well as the dialogue that occurs with each proposed implementations submitted as responses to the RFP. This dialogue forms an integral part of the overall SSF process and is essential to the production of semantically robust and relevant services for deployment in the clinical care, clinical research, and life sciences domains.

Alternatively, the SFM may be independently analyzed and one or more designs and implementations built independent of any OMG involvement (a path which appears likely in the case of the CRFQ SFM), a process not fundamentally dissimilar to the IHE's issuing of Profiles which vendors may choose to design/implement/test software around.

It is important to note that the SFMs developed under HL7 process specify the *functional requirements* of a particular service. In contrast, the OMG RFPs specify the *technical requirements* of a service. Finally, the resulting STM defines the resulting *technical model*, i.e. the design/deployment architecture (except as specified below). It should also be noted that in many cases, SFMs describe an overall coherent set of functional capabilities and / or define a minimum set of behaviors necessary to guarantee a minimal level-of-service when the service is deployed. A particular implementation (via its STM) may wish to *specialize* or *subdivide* these defined capabilities from either or both a *functional* or an *informational* (semantic) perspective, thereby defining (and providing) *de facto 'conformance profiles'* that may be used as the basis for the OMG RFP process and/or subsequent implementations.

1.1.2 Service Definition Principles

The high level principles regarding service definition that have been adopted by the Healthcare Services Specification Project are as follows:

- Service Specifications shall be well-defined and clearly scoped and be based on well-understood requirements and responsibilities;
- Services should have a unity of purpose (e.g., fulfilling one domain or area) (NOTE: this does *not* mean that services cannot/should not/will not be reused in across multiple domains and as part of multiple, potentially very different process which are constructed by orchestrating/choreographing multiple services. Rather, it means that the specific Interfaces and Operations of a given Service should be focused on a single service of coherent business purpose and value;
- Services may be utilized (i.e. 'orchestrated' or 'choreographed') with other services, thereby allowing their use in multiple contexts. (NOTE: changes in service context may require different semantic bindings).
- Services will be sufficiently specified so as to unambiguously address functional, syntactic, and semantic interoperability.
- It must be possible to replace one conformant service implementation with another that meets the same service specification without disrupting the overall functionality of a system providing the service.

With respect to the last bullet (above), a Service as specified by an SFM is regarded as a system component; the meaning of the term 'system component' in this context is consistent with Unified Modeling Language (UML) usage[1], i.e. a component is a modular unit with well-defined interfaces that is replaceable within its environment. A component can always be considered an autonomous unit within a system or subsystem. It has one or more provided and/or required interfaces, while its' internal details (i.e. its' implementation specifics) are hidden from clients and therefore inaccessible other than as provided by its interfaces.

Each SFM defines the interfaces that the service exposes to its environment, as well as the service's dependencies on services provided by other components in its environment. Dependencies in the Functional Model relate to services that have, or may in future have, a Functional Model at a similar level; detail dependencies on low-level utility services are not be included in an SFM and are considered out-of-scope for the SFM.

The manner in which services and interfaces are deployed, discovered, etc. also considered to be out-of-scope for an SFM. However, SFMs may reference content from other areas of HSSP work that deals with architecture, deployment, naming, etc. *Except where explicitly specified, these references are to be considered informative only.* (Note that an obvious exception to this statement is the case where one or more references are

[1] It is expected that services will be defined, in response to the OMG RFP process, as UML components, however that level of design is outside the scope of the Functional Model.

made to other SFMs relative to specific interface descriptions, e.g. an interface that is governed by an existing standard.) All other interactions within the scope of the scenarios identified above are potentially (depending on the service specifics) in the scope of a given SFM.

1.2 Overall Caveats and Disclaimers

- Examples included within a SFM are *illustrative* and *not normative* unless explicitly stated as such;
- The scope of information content of HSSP SFMs is not limited to HL7 content models (e.g. the Version 3 Reference Information Model). At a minimum, however, specifications should provide a semantic profile as part of its conformance profile to provide support for HL7 content models where applicable.
- ***The CRFQ SFM, the HL7 RIM is not directly referenced. Instead, the BRIDG Model – a domain analysis model adopted by the RCRIM TC as descriptive of the semantics of the RCRIM domain – is the reference information model utilized by the CRFQ SFM. An element-by-element mapping of the BRIDG Model – in which each attribute is bound to a formal V3 Abstract Data Type, identical to the attribute-binding approach of the RIM -- to the RIM is available at www.bridgmodel.org. The presence of this mapping means that automated transforms between BRIDG and RIM representations of a given set of semantics can be developed if needed.***

1.3 Context of this SFM within the HSSP Roadmap

The CRFQ service forms an important component within the larger context of service specifications because it is the first SFM defined in the domain of clinical research/life sciences. However, because the CRFQ service is expected to exist within larger business contexts (e.g. the pharmaceutical context, clinical research organizations, etc.), it is expected that the CRFQ service will effectively leverage the earlier HSSP specifications with respect to both content and functionality. In particular, the CRFQ service is expected to have dependencies on one or more infrastructure services as suggested in the following diagram:

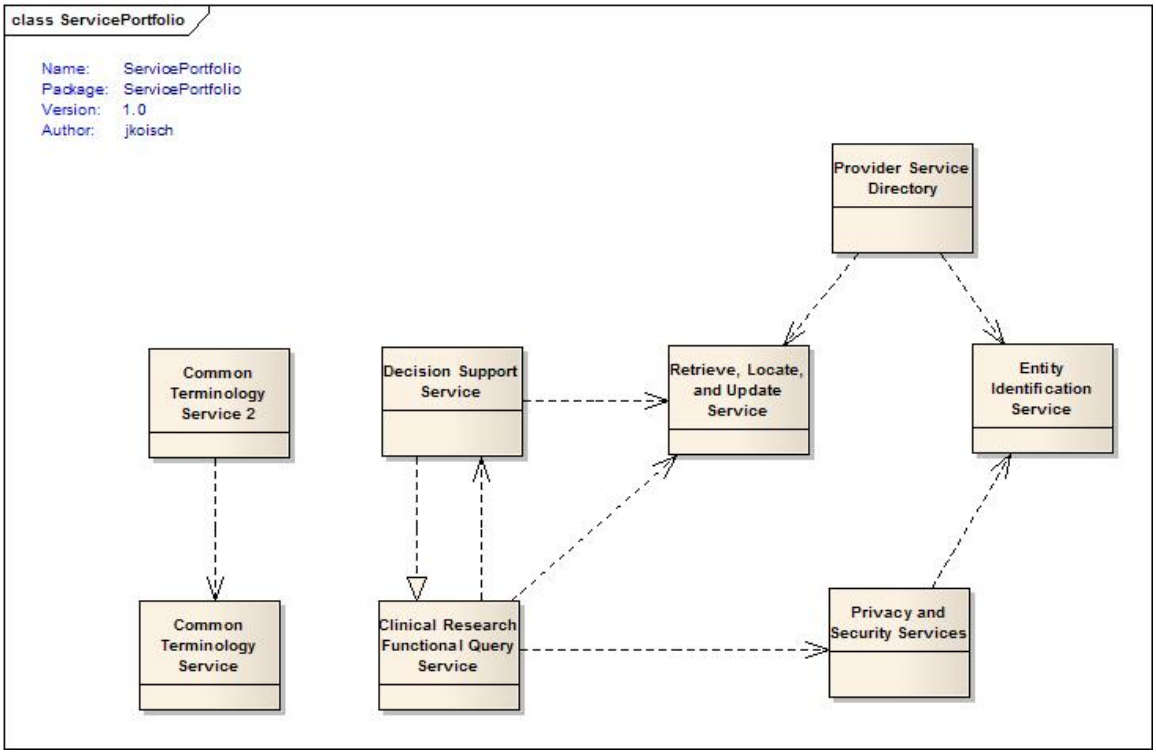


Figure 1: The potential relation of CRFQ to other HSSP Services

The Clinical Research Filtered Query (CRFQ) service is one of several 'business-level' services which are anticipated to utilize other HSSP 'infrastructure' and 'process' services (e.g., RLUS, DSS, Consent, Anonymization, etc.) in business-process-driven orchestrations designed to achieve critical business goals in an automated/semi-automated fashion.

2 CRFQ Service Overview and Business case

2.1 CRFQ Service Description and Purpose

As described above, the purpose of an HSSP SFM is to identify and document the functional requirements of services important to healthcare (including, in the case of the CRFQ service, clinical research and its intersection with healthcare). Accordingly, this SFM defines the functional requirements of a Clinical Research Filtered Query (CRFQ) service, a service which provides a set of capabilities in the context of clinical trials, protocols and associated protocol metadata. In particular, the CRFQ service focuses on the relationship between individual person/animal/etc. genotypic/phenotypic data and the so-called 'Inclusion and Exclusion (I/E) (or Eligibility) criteria' associated with a protocol, i.e. the characteristics that are considered essential as being present (or absent) in a person/animal/etc. in order for that person/animal/etc. to be viewed in the context of a particular protocol as a potential subject.

The CRFQ service is defined to exist in two types of contexts: as a service on a protocol repository which filters individual protocols – based on protocol meta-data describing the protocol's Inclusion and Exclusion criteria – against incoming *individual* person/animal/etc. to find one or more protocols in which the person/animal/etc. may qualify as a research subject; and as a service on an EHR repository which filters individual protocol meta-data (Inclusion and Exclusion criteria) against patient data to find a suitable potential 'cohort' for the protocol. In addition, if one generalizes the latter notion of 'finding cohorts based on defined signal descriptions as inputs,' the CRFQ service may also be used in the context of 'real-time safety monitoring,' i.e. the desire to search an EHR repository for a set of patients satisfying a particular 'signal definition.'

It should be noted that the CRFQ SFM is specifically restricted to application in the Clinical Trials context. It should not be seen as a generalized query service, but rather was initially scoped to the clinical research domain because of a clearly-defined business need. However, as mentioned above, other context may find its basic structure readily accessible and extensible and may therefore choose to use it as the basis for the development of more generalized query services (including the identification of additional semantic profiles other than those listed in the context of the CRFQ.) In any one of these situations, implementers may find the need to modify interface and/or operation names given that the current names were specifically chosen because they represent well-known concepts in the clinical research domain.

CRFQ will provide a foundational component for other services as both service consumer and provider, including DSS. Additionally, it is expected to be a significant motivation for the adoption of standards in both the Pharma/Clinical Trial and Healthcare domains as both parties recognize the value in having standard semantic profiles that can bind a set of standard interfaces such as those defined by the CRFQ service.

2.2 Scope

The previous section presented a high-level outline of the CRFQ service. Of particular importance is the fact that the service is scoped to cover differing -- but related -- aspects of the *domain of clinical research*, i.e. the domain defined as 'protocol-driven research involving human subjects.' As a consequence, the goal of the application of the CRFQ service is the efficient pairing of potential subjects with either protocols in the context of exchange/comparison of computable demographic, phenotypic, and/or genotypic I/E criteria associated with both protocols and potential subjects (e.g., patients), as well as the related domain of real-time safety monitoring when the safety events-of-interest have been sufficiently well defined to be syntactically and semantically similar to protocol inclusion or exclusion criteria.

Successful deployment of the CRFQ service assumes the following:

- the presence of protocols encoded with protocol-specific metadata that sufficiently describe the protocol's I/E criteria; and
- the existence of patient repositories with sufficient amounts of appropriately encoded demographic, phenotypic, and/or genotypic data to enable automated comparisons of patient data with protocol I/E criteria.

The CRFQ's composite functionality presents these facets of existing (often-by-hand) clinical research queries to meet the different business needs of pharma and clinical research organizations in addition to patients (and/or their providers) searching for possible protocols in which they might participate.

Specifically excluded from – but, in many cases, essential to the successful and relevant application of – the CRFQ service are additional, closely-coupled functions (which could be provided as services – see Section 1) including:

- Security services including identification, authentication, and authorization (S services)
- Resource Location and Update service (RLU service)
- Decision Support/Inference service (DS service)
- Consent management service (Trusted Broker) Service (CM/TB service)
- Anonymization/pseudonymization service (A/P service)

CRFQ will provide a foundational component for other services as both service consumer and provider, including DSS. In particular, with respect to the difference between the CRFQ service and the Decision Support Service (DSS) which has previously been developed under the auspices of the Service Specification Framework (SSF), it should be noted that the CRFQ service is concerned with relatively coarsely granulated, often atomic data elements (e.g. age, gender, lab data, disease-specific severity classifications, etc.) whereas the DSS is designed to perform inferences on data, some of which could

conceivably come from the results of an invocation of the CRFQ service. A *possible* scenario (*included for exemplar purposes only)* involving the application/orchestration of all of the services listed above is:

Client business scenario → Security → CRFQ → Resource Location → Consent → (Anonymization) → Decision Support → Return to Client

Details of this service orchestration are beyond the scope of this document, but are included here for correlation by readers who are familiar with both the domains of clinical research trials and decision support.

2.3 The Rationale for CRFQ

The CRFQ service is driven by business needs in the clinical research domain as manifested in four distinct business use cases:

- *Scenario #1: Patients/Providers Searching for Protocols* – Individual patients (or their providers) may submit individual demographic, phenotypic, and/or genotypic data against a repository of protocols, the specific I/E criteria of which are available in a computable form to determine which protocols the patient could be eligible for should they so choose to participate.

- *Scenario #2: Protocols Searching for Potential Subjects* – For a specific protocol and its computable I/E criteria, identify a cohort of potential trial subjects from a repository of patient data by comparing the protocol's I/E criteria with individual patient demographic, phenotypic, and/or genotypic characteristics. (NOTE: this scenario may also have applicability in other related domains such as Quality Measurement/Assessment, etc.)

- *Scenario #3: Sponsors Evaluating I/E Criteria Efficacy* -- For a given single (or small set of) proposed I/E criteria, query one or more sources of patient data to determine the sensitivity or effectiveness of a particular I/E criterion to ensure that when the protocol is actually published, a suitable subject base will be able to be identified.

- *Scenario #4: Monitoring Patient Populations for known Safety-related events-of-interest ('signals')* – From a functionality perspective, this scenario – in addition to be essentially analogous to concurrent adverse event monitory and reporting as

is practiced by the pharamco-vigilence processes -- is virtually identical to scenario #2 (above), the only difference being the business context in which it is applied, In this context, the 'I/E' criteria that are applied against a patient repository are the I/E criteria that define the 'safety-related event-of-interest,' the list of patients being returned therefore being the list of patients who have potentially experienced a safety-related event-of-interest. (**NOTE: Scenario #4 is *not further* illustrated in this document (and will only be occasionally mentioned following this discussion) because of its virtually identical functional and informational content to scenario #2.)**

The overall business context of the CRFQ service for scenarios # 2, #3, and #4 is shown in the following graphic with the orange circle. CRFQ represents a service-oriented approach to providing specific, consistent functionality across deployments. When combined with common information models (BRIDG), CRFQ provides the potential for a consistent mechanism for implementing eligibility filtering within and among organizations.

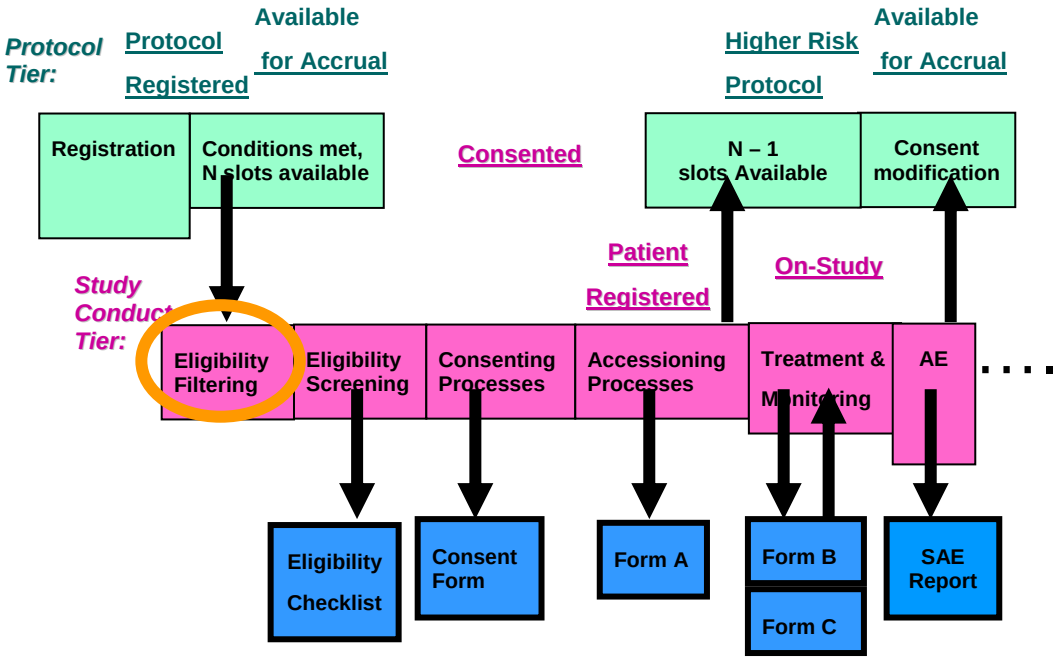


Figure courtesy of Joyce Niland, City of Hope

Viewed from a slightly different perspective, the general flow of protocol design from inception to subject recruitment in both the drug-development and clinical research contexts – and therefore focused on the application context of Scenarios #2 and #3 – is depicted below (note the additional step in the Pharma flow of identifying PIs before recruiting subjects. Additional process steps the distinguish single- from multi-site trials are specifically omitted as beyond the scope of the current SFM):

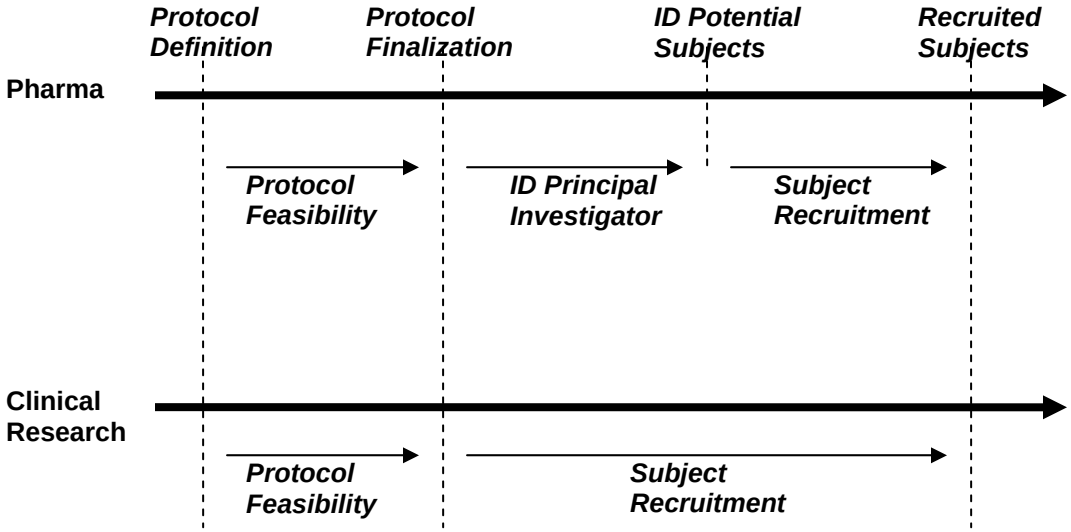


Figure courtesy of Isabelle de Zegher, Novartis

2.4 Structure of the CRFQ Service

As will be discussed in detail in subsequent sections of this document, one can view the CRFQ service as a function (in the mathematical sense of the term) that has a set of 'input parameters' and returns a value based on these inputs. In particular, the four previously-listed Scenarios above define the function(s):

F_1 and F_2 with the following characteristics:

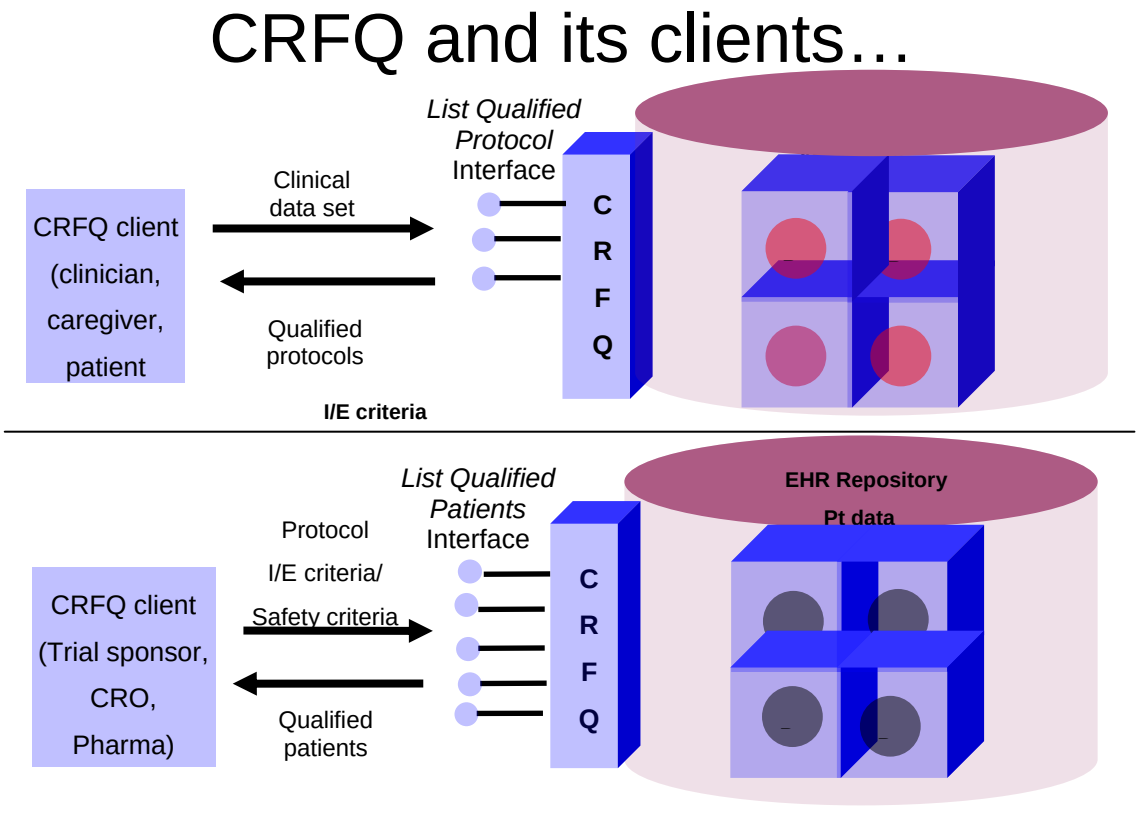
F_1 (some input) → list of qualified protocols (Scenario 1)

F_2 (some input) → list of qualified patients (Scenarios 2, 3, 4)

When one digs a bit deeper for the specifics of 'some input' for each of the functions, one finds that the input data for the two functions are *identical* in content, but utilized differently by the two functions. In particular, both of the functions listed above require the following as input data sets *although it should be noted that these may be either explicit or implicit on the interface:*

- Inclusion criteria (associated with a protocol)
- Exclusion criteria (associated with a patient)
- Patient-specific data

In the case of F_1, the patient-specific data is used to filter protocols whose inclusion and/or exclusion criteria allow the patient to consider participating in the protocol. In the case of F_2, the protocol's inclusion and/or exclusion criteria are used to filter patients whose specific data will allow them to meet (or not be excluded) by those criteria. Thus, the CRFQ service's overarching *business* functionality, i.e. the context in which it is applied, is, in fact, a 'usage context' that is defined – at a high level – by which way the CRFQ service is 'facing,' i.e. outward to its clients from a patient repository, or outward to its clients from a protocol repository. These two client relationships are shown in the following figure:



The results of a call to the CRFQ service will thus be either a list of potential protocols or a list (ranging in detail from a simple count of list members to detailed information about each list member) of potential subjects. Various business process/service coordination details that occur in the course of a CRFQ invocation, e.g. the concurrent orchestration of the CRFQ service with security services, infrastructure services (e.g. RLUS), consent, management services, decision-support services, and/or anonymization / pseudonymization services are discussed elsewhere this document, but are generally out of scope for the purposes of specifying the behavioral interface.

From an IT vendor viewpoint, CRFQ represents a core business capability that could be included in either EHR applications (which are increasingly living at the intersection of Clinical Research/Clinical Trials and healthcare) or clinical trial applications. In addition, Personal Health Record (PHR) applications could support the role of their application as a client of the *List Qualified Protocols* interface of CRFQ.

Regardless of how the service is established for particular deployment contexts, CRFQ represents a powerful tool for researchers and patients wishing to participate in clinical trials. By providing a functionally consistent set of interfaces, a subset of business-oriented system behavior is specified. This behavior supports the two core scenarios currently pervasive within the clinical trial community: finding protocols for a particular patient, and finding a patient population for a particular protocol. Support of CRFQ means that researchers could have access to broader populations on which to conduct their trials, while patients might have access to a broader, more diverse set of trials in which they could participate.

Intrinsic in both of these scenarios in a non-ambiguous definition of the information content exchanged by the service client and service provider, i.e. the 'semantics bound to the messages passed during service invocation.' Considerable progress has been made in standardizing clinical trial and healthcare semantics (e.g the BRIDG Model and the HL7 Reference Information Model). However, business-level interactions between systems based on standardized interfaces are still relatively rare. The HSSP process in general, and the SFM in particular provide a strategy to accomplish loose coupling informational components to functionally consistent components, In this context, CRFQ represents a delivery mechanism for best-of-breed information models that support clinical trial research without being hamstrung by potential obsolescence.

Should a more generalized *Functional Query* service (not bound to the functional semantics of clinical research) emerge in the future, the migration from CRFQ should take into account the following items:

- The naming of operations and parameters are informed by the clinical research domain
- CRFQ posits that a protocol or patient repository is a dependency of the service interface. Should a more abstract notion of functional query interface be conceived, a similarly abstract notion of repository dependency should accompany it

- CRFQ might be called a business level service in that it would likely sit closer to the end user interface (or client) in the service stack. Because of this, it represents a level of granularity that is appropriate to the business of querying data to support clinical research. Other sorts of queries, or indeed businesses, may require different level of granularity.

2.4.1 The Two CRFQ Interfaces: Common Input Parameters, Different Functions

As noted above, the functions F_1 and F_2 (which are the formal representations of the two interfaces of the CRFQ service) share a common set of inputs. Closer examination of the two functions, however, reveals that there are differences in the character of the inputs. In particular, F_1 (List Qualified Protocols) takes a set of data from a single patient and returns an array of protocol IDs, names, etc. (Note that this framework is similar to that in which a set of patient data is aligned to one or more Quality Measures. As noted above, there are a number of contexts in which CRFQ may be applied). In contrast, F_2 (List Qualified Patients) takes a set of the metadata from a single protocol and returns an array of patient IDs (which may be pseudo-IDs depending on security considerations) (Again, note the similarity to the Public Health context and Quality Measurements being used as 'Eligibility Criteria' in the filtering of patient data.) There are other differences in the input data as well (e.g. F_2 allows optional patient preferences and scoring/weighting criteria whereas $F_{1\ does}$ not, etc.) which indicate that there are, in fact, two distinct functions represented by the two interfaces of the CRFQ Service definition. Regardless, these differences are quantitative rather than qualitative. For that reason, the operations themselves are profiled in Section 6 to highlight these functional differences.

2.4.2 Representative Examples of Deployment Scenarios

Scenario #1: A 39 year old woman recently diagnosed with Stage III breast cancer goes on the Web to search for potential experimental treatment. A query shows that there are 58 studies that encompass Stage III breast cancer in the eligibility. The woman is otherwise healthy, with no prior cancers, surgeries, chemotherapy or hormone therapy, and no evidence of cardiovascular, kidney, lung, or neurological disease, so all of these trials still remain open to her when she enters this health information. The woman then reviews her recent test results performed by her local physician, and enters into the search engine that she is Estrogen Receptor positive (ER +), Progesterone Receptor negative (PR -), and HER2 Nu -. This pattern reduces the number of available trials from 58 to 6 for which she may be qualified based on further screening of the complete eligibility criteria for this trial.

Scenario #2: A pharmaceutical study for Stage I-II breast cancer requires women who are ages 50+ and who are post menopausal, and have a good performance status and organ functioning. The women cannot be prior smokers, and they can have received

prior chemo or hormone therapy, as long as it ended more than 6 months ago. Prior mastectomy or breast conserving surgery are not allowed. Applying these filter criteria to an electronic health record database of ~10,000 in an oncology practice yields 145 women who may be eligible for this trial, pending searching of further detailed eligibility criteria.

Usage of CRFQ starts from the perspective of four business contexts spanning two interfaces:

INTERFACE: *List Qualified Protocols*

- *Scenario #1: Patients/Providers searching for Protocols*

INTERFACE: *List Qualified Patients*

- *Scenario #2: Protocols searching for Potential Subjects*
- *Scenario #3: Sponsors evaluating I/E Criteria Efficacy*
- *Scenario #4: Monitoring patient populations for known safety-related events-of-interest ('signals')*

As noted earlier, the character of the parameters for these functions, while qualitatively the same, differ in structure or content. *List Qualified Patients*, for example, can list (output of the function) those patient by name, by ID, or by weighted criteria, depending on the context.

2.5 Existing Semantic Models from the Clinical Trial Industry/Context

Three semantic models exist in the world of clinical trials and are intended to be used within the context of CRFQ:

- The BRIDG Model
- The ASPIRE data set
- HL7 Version 3 Data Type Specification

Additional semantic content can be expected to be defined under each of these (as well as possibly other) projects over the course of the next 6-12 months and would be expected to be included in the CRFQ service RFP in the form of template semantics.

These models have both a successful history and wide acceptance within the larger clinical trial community.

BRIDG Model:

The existence of the BRIDG Model (http://www.bridgmodel.org), a Domain Analysis Model (DAM) whose scope is "protocol-driven research involving human, animal, or device subjects and all associated regulatory artifacts' has resulted in the CRFQ SFM

development team deciding that the service defined in this document should be semantically bound to the BRIDG Model rather than the HL7 RIM per se. The semantics of the BRIDG Model are, in fact, mappable to the RIM. As a result, the binding of the CRFQ service to BRIDG rather than to the RIM is more a statement of the need for clearly stated domain semantics (which are present in the BRIDG Model by virtue of it's being a DAM focused on a specific domain which, not coincidently, is the domain in which the CRFQ service will be deployed) than a departure from RIM semantics. As noted above, an element-by-element mapping of BRIDG Model attributes to RIM elements is available at www.bridgmodel.org.

The stakeholders in the CRFQ team extend outside of the traditional bounds of HL7 participation, but are encompassed by the current list of BRIDG stakeholders. Specifically, these include:

- HL7 RCRIM
- CDISC (representing pharma)
- NCI
- FDA

It is the collective desire of these stakeholders and the other members of the CRFQ team not directly linked to these stakeholder organizations to have the SFM's semantic profiles linked to the domain-specific BRIDG Model (or a semantically equivalent RIM representation) rather than the cross-domain HL7 Reference Information Model.

ASPIRE Project:

NOTE: The ASPIRE Project is described in the CRFQ SFM as an example of an effort to codify and standardize protocol metadata. It is *not at present*, however, a vetted, balloted standard of any organization. Nor should readers of this document mistakenly conclude that the assignment of the CRFQ SFM as an ANSI DSTU implies in any way that an equivalent designation is being *de facto* granted to the ASPIRE data set. *To repeat: The following discussion/documentation on the ASPIRE Project and its data are included in this document purely as an exemplar semantic profile for CRFQ.*

The overall objective of ASPIRE is to create a structured representation of a core set of encoded protocol eligibility criteria, using accepted medical terminology and vocabulary standards when available. The goal is to strike a practical middle ground of core eligibility criteria that it would be feasible for all trials to code against, rather than to develop complete coded eligibility criteria, which would be a much more extensive and separate project. Through uniform coding of such a core set of eligibility criteria it will be possible to facilitate more rapid efficient screening of potential participants for available clinical trials, potentially worldwide, thereby speeding the discovery of new interventions to treat, prevent or screen for disease among patients. Coded eligibility criteria also will provide semantic interoperability (the exchange of content and meaning) among systems and stakeholders, serve as the underpinning for various technical implementations for subject screening and recruitment, facilitate electronic protocol

authoring, and standardize the core eligibility components of protocol registration across all trials. The ASPIRE project is proceeding in two phases:

1. Recommendations for common semantics for coded eligibility criteria across all disease types ("pan-disease" eligibility criteria) and
2. Recommendations for common semantics for coded eligibility criteria within several major disease types (disease-specific eligibility criteria)

Included in the above work are the following specific tasks:

- Specification of use cases for eligibility encoding and protocol 'filtering' to automatically identify potential studies suitable for patients and patients potentially eligible for open trials
    - Phase 1: Creation of a spreadsheet for coded terms for pan-disease core eligibility criteria, leading to a metadata structure for the required data dictionary
    - Phase 2: Creation of a spreadsheet for coded terms for several selected disease-specific core eligibility criteria, leading to a metadata structure for the required data dictionary
- Specification of harmonizable artifacts to be incorporated into BRIDG
- Proposal for the business process for continuing the evolution of additional core criteria and coverage of more diseases, including updating, review, and approval of proposed data elements
- Establishing a mechanism for measuring successful utilization and technical implementation of encoded eligibility criteria under several use cases

In 2006 the project subgroup met over several months to establish the mission and charter for the ASPIRE subproject, establish the protocol filtering use cases driving the vocabulary analysis, and conduct the analysis for Phase I, for the pan-disease core coded eligibility criteria. A proposed set of Phase I eligibility criteria was completed in 2006, for subsequent vetting with the CDISC PR group and other interested parties, such as RCRIM and the CDASH initiative. In 2007 the proposed pan-disease core eligibility criteria were established and vetted with the CDISC Protocol Representation group, and disease-specific criteria for 2 disease areas, breast cancer and diabetes. The next disease to be evaluated for common eligibility criteria will be pediatric hypertension, which will be the subject of an FDA pilot for evaluating coded new product submissions.

The ASPIRE project is focused on defining protocol metadata that may be computationally compared to individual patient data to determine whether a given patient might be eligible for a given protocol. The project has separated protocol metadata into two types:

- **Pan-disease data** (e.g. age, gender, etc.). To date, the ASPIRE project has defined pan-disease data elements.
- **Disease-specific data.** To date, the ASPIRE project has defined data for Breast Cancer and Diabetes

631 The tables below depict the pan-disease and disease specific data:

632 **Pan Disease Criteria**

Element Name	Definition	Codes	Attributes	Data Type	Notes (examples from existing protocol repositories such as CT.gov and PDQ (NCI), etc.)[2]
Demographic Criteria					
Min Age	Minimum allowable age at entry into study		Units of measure: Hours Days Weeks Months Years	Num	CT.gov: Minimum _age PDQ: Low/ Age Codes harmonized with SDTM
Max Age	Maximum allowable age at entry into study		Units of measure: Hours Days Weeks Months Years	Num	CT.gov: Maximum _age PDQ: High/Age Codes harmonized with SDTM
Sexpop (Gender)	Allowable gender(s) on study	Male Female Both	NA	Char	Harmonized with SDTM
TINDTP (Primary Purpose)	Primary purpose for conducting the study	Prevention Screening/Detection Treatment Symptom Management Quality of Life		Char	Use to Branch to intent specific Inclusion/Exclusion Criteria Attributes/Element name harmonized with SDTM

[2] www.clinicaltrials.gov and www.cancer.gov/cancertopics/pdq

Element Name	Definition	Codes	Attributes	Data Type	Notes (examples from existing protocol repositories such as CT.gov and PDQ (NCI), etc.)[2]
Type	Type of trial	Safety Efficacy Bio-Equivalence Bio-Availability Confirmation Exploratory Pharmacoecomnomic Pharmacogenomics Pharmacokinetics Pharmacodynamics	NA	Char	
Targeted Hispanic Ethnicity	Targeted inclusion of Latinos/Hispanics	Yes/No	NA	Char	
Targeted Minority Racial Groups	Targeted inclusion of minorities based on race	American Indian or Alaska NativeAsianBlack or African AmericanNative Hawaiian or Other Pacific Islander	NA	Char	Code all that are specifically targeted
Perf Status	Level(s) of function included on study	*Able to carry on normal activity and to work; no special care needed *Unable to work; able to live at home and care for most personal needs; varying amount of assistance needed *Unable to care for self; requires equivalent of institutional or hospital care; disease may be progressing rapidly	NA	Char	Can be derived from KPS, ECOG, Lansky

Element Name	Definition	Codes	Attributes	Data Type	Notes (examples from existing protocol repositories such as CT.gov and PDQ (NCI), etc.)[2]
Repro Status of subject	Allowable reproductive status	Active Inactive Either	NA	Char	Applies to M and F
Repro Status of partner	Allowable reproductive status	Active Inactive Either	NA	Char	Applies to M and F
Pregnancy	Allowable status with respect to pregnancy	Pregnant Not Pregnant Either	NA	Char	Applies to F
Nursing Status	Allowable status with respective to nursing	Active Inactive Either	NA	Char	Applies only to F
Diagnostic Criteria					
TDIGRP (Diagnosis Group)	Diagnosis Group	Healthy ICD-9 ICD-10	Healthy Subjects	Char	Definition and Codes harmonized with SDTM
Other Health Conditions					
History of Cardiac Disease	Allowable status with respect to having prior cardiac disease history	AllowedAllowed with ConditionsExcludeRequiredExclude with ConditionsNot Mentioned	NA	Other	NA
History of Kidney Disease	Allowable status with respect to having prior kidney disease history	Allowed Allowed with Conditions Exclude Required Exclude with Conditions Not Mentioned	NA	Other	NA
History of Lung Disease	Allowable status with respect to having prior lung disease history	Allowed Allowed with Conditions Exclude Required Exclude with Conditions Not Mentioned	NA	Other	NA

Element Name	Definition	Codes	Attributes	Data Type	Notes (examples from existing protocol repositories such as CT.gov and PDQ (NCI), etc.)[2]
History of Liver Disease	Allowable status with respect to having prior liver disease	Allowed Allowed with Conditions Exclude Required Exclude with Conditions Not Mentioned	NA	Other	NA
History of Neurologica l Disease	Allowable status with respect to having prior neurological disease	Allowed Allowed with Conditions Exclude Required Exclude with Conditions Not Mentioned	NA	Other	NA
History of smoking	Allowable status with respect to having prior smoking history	AllowedAllowed with ConditionsExclu deRequiredExclu de with ConditionsNot Mentioned	NA	Other	NA
Current Smoking	Allowable status with respect to having prior smoking history	Allowed Allowed with Conditions Exclude Required Exclude with Conditions Not Mentioned	NA	Other	NA

633

634 **Breast Cancer Criteria**

Element Name	Definition	Codes	Attributes	Data Type	Notes
Diagnostic Criteria					
Current Stage	Stage required for study participation	DCIS I II III IV I - III Recurrent N/A		Char	Check all that apply

Element Name	Definition	Codes	Attributes	Data Type	Notes
Tumor Size	Maximum allowable tumor size Minimum allowable tumor size	NA	size in cm	Num	
Other Health Conditions					
Active Brain Metastases	Brain metastases that have not been treated or are not responding to treatment	Allowed Allowed with Conditions Exclude Required Exclude with Conditions Not Mentioned	NA	Char	
Estrogen Receptor Status	Inclusion based on patient's estrogen receptor status	Positive Positive with conditions Negative Known N/A	NA	Char	
HER2/neu Receptor Status	Inclusion based on patient's HER2/nu receptor status	Positive Positive with conditions Negative Known N/A	NA	Char	
Prior malignancy	History of malignancy other than breast cancer	Allowed Allowed with Conditions Exclude Required Exclude with Conditions Not Mentioned	NA	Char	
Progesterone Receptor Status	Inclusion based on patient's progesterone receptor status	Positive Positive with conditions Negative Known N/A	NA	Char	
Prior Treatment					

Element Name	Definition	Codes	Attributes	Data Type	Notes
Prior Chemotherapy	Inclusion/exclusion based upon prior chemotherapy	Allowed Allowed with Conditions Exclude Required Exclude with Conditions Not Mentioned			
Breast Conservation Surgery (Lumpectomy)	Inclusion/exclusion based upon prior breast conservation surgery	Allowed Allowed with Conditions Exclude Required Exclude with Conditions Not Mentioned			
Mastectomy	Inclusion/exclusion based upon prior mastectomy	Allowed Allowed with Conditions Exclude Required Exclude with Conditions Not Mentioned			
Prior Endocrine/Hormone Therapy	Inclusion/exclusion based upon prior hormone therapy	Allowed Allowed with Conditions Exclude Required Exclude with Conditions Not Mentioned	NA	Char	

635

636 **Diabetes Criteria**

Element Name	Definition	Codes	Attributes	Data Type	Notes
Diagnostic Criteria					
Type of Diabetes Diagnosis	Diabetes type required for study participation	Type I, Type II, Either	NA	Char	
Minimum duration of diabetes	The minimum duration since diabetes first diagnosed	Allowed Allowed with Conditions Exclude Required Exclude with Conditions Not Mentioned	Years	Num	

Element Name	Definition	Codes	Attributes	Data Type	Notes
Minimum duration of stable diabetes	The minimum duration patient has been stable	Allowed Allowed with Conditions Exclude Required Exclude with Conditions Not Mentioned	Months	Num	Added after teleconference. Expert said typically minimum 1 month for treatment naïve, 3 months for concomitant metformin
Other Health Conditions					
History of Gestational Diabetes	Any prior diagnosis of gestational diabetes	Allow with conditions Not allowed Required Not required N/A in protocol	NA	Char	Yes/No
History of Pancreatic Cancer	Any prior diagnosis of pancreatic cancer	Allow unconditionally Allow with conditions Not allowed Required Not required N/A in protocol	NA	Char	Yes/No
History of Pancreatitis	Any prior diagnosis of pancreatitis	Allow unconditionally Allow with conditions Not allowed Required Not required N/A in protocol	NA	Char	Yes/No
Maximum baseline HbA1c	Maximum glycosylated hemoglobin level at baseline (e.g., cutpoint < 9.5 to 11)	AllowedAllowed with ConditionsExcludeRequiredExclude with ConditionsNot Mentioned	%	Num	
Maximum body mass index (BMI)	Maximum BMI at baseline	Allowed Allowed with Conditions Exclude Required Exclude with Conditions Not Mentioned	kg/m2	Num	

Element Name	Definition	Codes	Attributes	Data Type	Notes
Maximum body weight	Maximum body weight at baseline	Allowed Allowed with Conditions Exclude Required Exclude with Conditions Not Mentioned	pounds or kg	Num	
Maximum creatinine clearance	Maximum creatinine clearance at baseline	Allowed Allowed with Conditions Exclude Required Exclude with Conditions Not Mentioned		Num	
Maximum duration of diabetes	The maximum duration since diabetes first diagnosed	AllowedAllowed with ConditionsExcludeRequiredExclude with ConditionsNot Mentioned	Years	Num	Expert said typically don't specify maximum so can delete
Minimum baseline HbA1c	Minimum glycosylated hemoglobin level at baseline (e.g., cutpoint >7 to 7.5)	Allowed Allowed with Conditions Exclude Required Exclude with Conditions Not Mentioned	%	Num	
Minimum body mass index (BMI)	Minimum BMI at baseline	Allowed Allowed with Conditions Exclude Required Exclude with Conditions Not Mentioned	kg/m2	Num	
Minimum body weight	Minimum body weight at baseline	Allowed Allowed with Conditions Exclude Required Exclude with Conditions Not Mentioned	pounds or kg	Num	

Element Name	Definition	Codes	Attributes	Data Type	Notes
Ratio of waist to hip circumference	Minimum ratio of waist to hip circumference at baseline	Allowed Allowed with Conditions Exclude Required Exclude with Conditions Not Mentioned	NA	Num	Is this an eligibility criteria? Or mainly a baseline and outcome measurement?
Serum creatinine	Maximum serum creatinine at baseline	Allowed Allowed with Conditions Exclude Required Exclude with Conditions Not Mentioned		Num	
Waist circumference	Minimum waist circumference at baseline	AllowedAllowed with ConditionsExcludeRequiredExclude with ConditionsNot Mentioned	inches or centimeters	Num	Is this an eligibility criteria? Or mainly a baseline and outcome measurement?
Prior Treatment					
Prior use of sulfonylurea	Any use prior to randomization	Allowed Allowed with Conditions Exclude Required Exclude with Conditions Not Mentioned	NA	Char	Yes/No with optional conditions--typically require about 3 months washout.
Prior use of exogenous insulin	Any use of insulin prior to randomization	Allow unconditionally Allow with conditions Not allowed Required Not required N/A in protocol	NA	Char	Yes/No with optional conditions
Prior use of thiazolidinediones (TZD)	Any use of TZDs prior to randomization	Allow unconditionally Allow with conditions Not allowed Required Not required N/A in protocol	NA	Char	Yes/No with optional conditions. Usually not allowed since very long washout period.

637

HL7 Version 3 Abstract Data Type Specification:

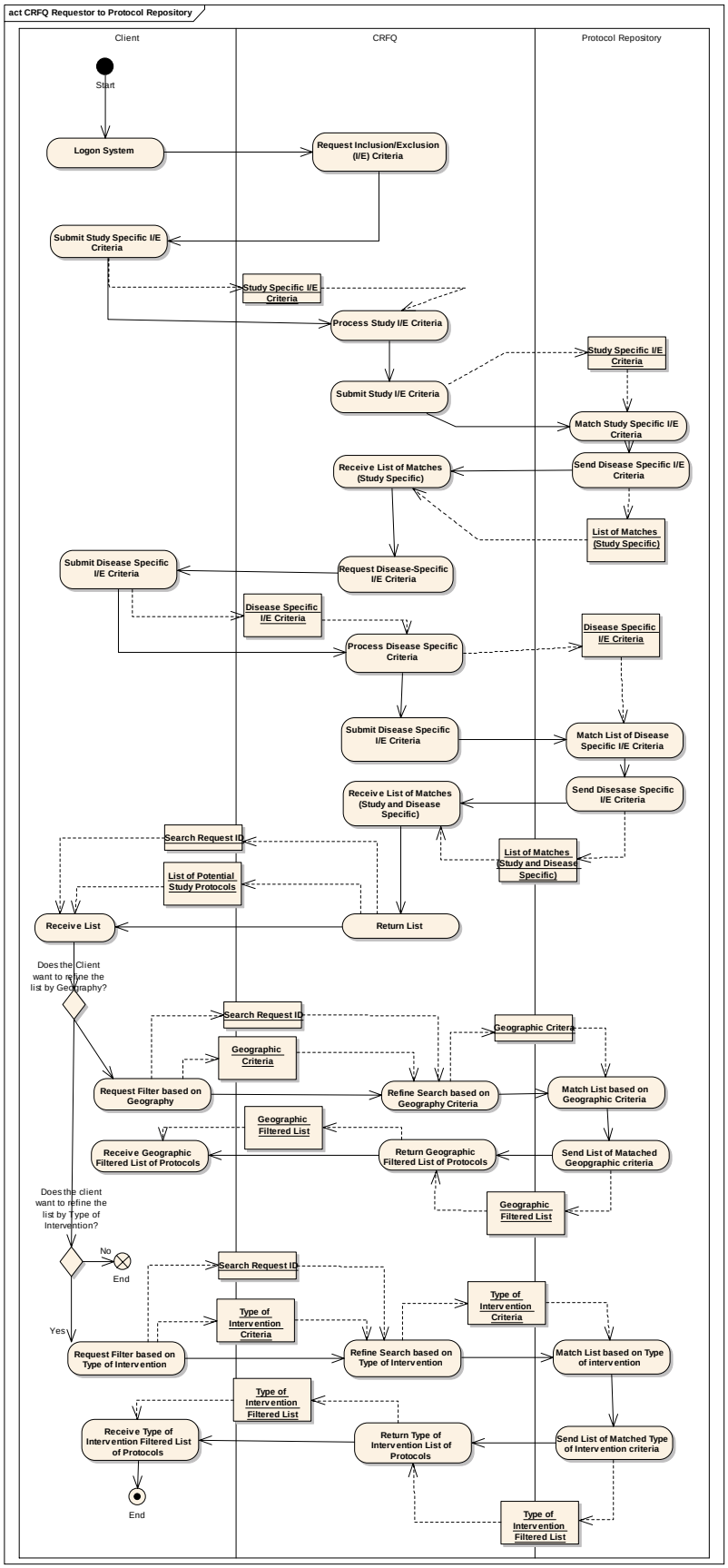
The HL7 Version 3 Abstract Data Type Specification Release 1 (http://www.hl7.org) is an American National Standards Institute (ANSI) **s**tandard which defines the semantics of a collection of complex data types (e.g. Physical Quantity, Coded Description, General Timing Specification, etc.) which HL7 has found to be essential if machines are going to exchange data at a computable semantically interoperable level. (NOTE: as of the voting on the CRFQ, Release 2 of the ADT specification was in final ballot. Future releases of the BRIDG Model will be, as will the HL7 Reference Information Model's future releases, bound to R2 rather than R1), it is beyond the scope of this SFM to describe the data types in detail. However, it should be noted that *each attribute* in the BRIDG Model is bound to an HL7 V3 data type specification and any implementation of the CRFQ service will be expected to support the necessary V3 data types required to express the I/E criteria-of-interest to the CRFQ client.

3 Business Scenarios

As mentioned in Section 2, the CRFQ service will expose two interfaces in support of four basic business scenarios. Exemplar storyboards and associated UML Activity Diagrams are presented here for Scenarios #1, #2, and #3. *As previously mentioned, Scenario #4 is not being discussed because it is essentially identical to Scenario #2 with the replacement of protocol-specific I/E criteria with safety-signal-specific events-of-interest.*

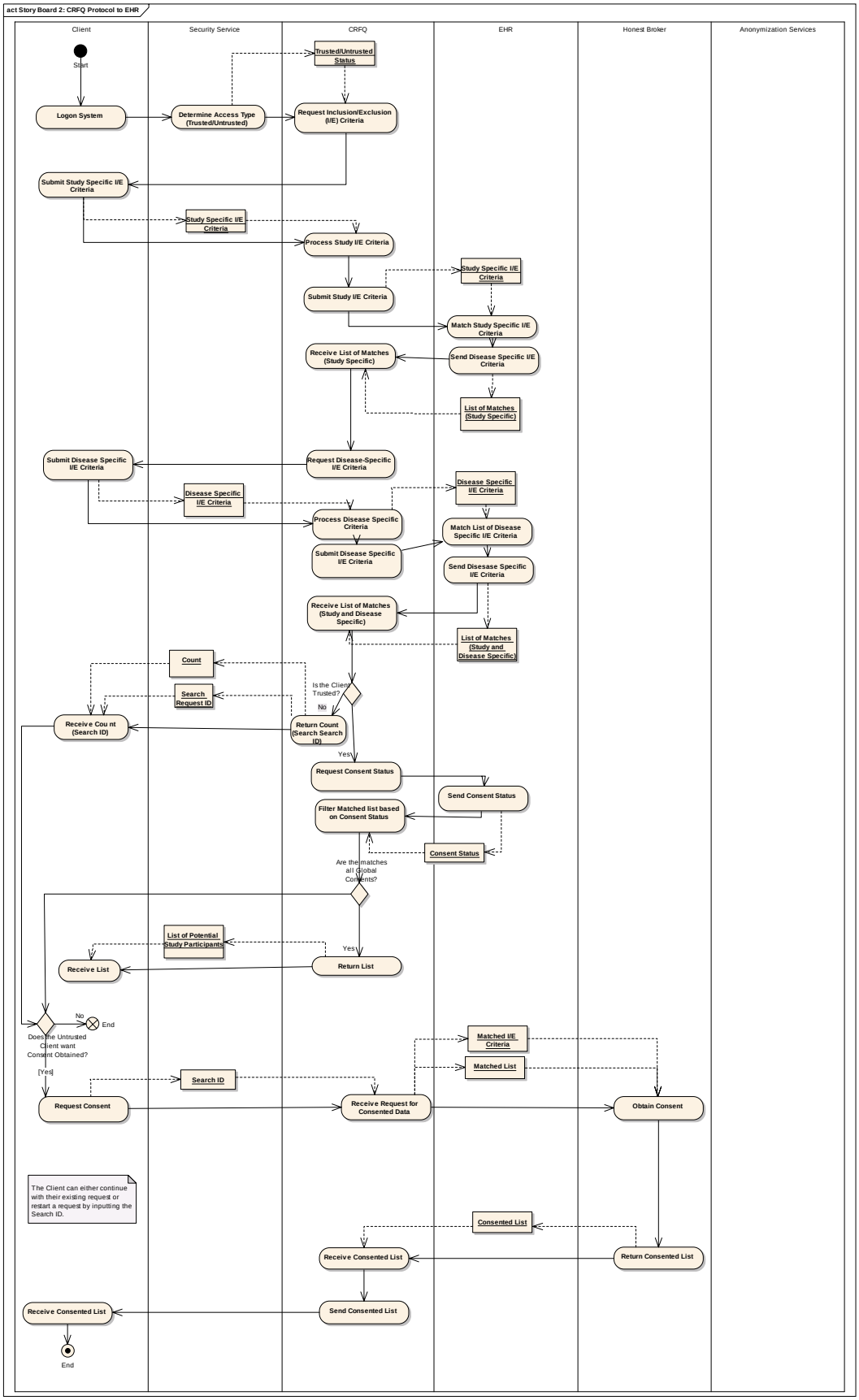
3.1 Scenario #1: Patients/Providers searching for Protocols

A woman (or her clinician or other caregiver acting on her behalf) with breast cancer is searching for possible leading-edge experimental treatment options. Using the CRFQ web interface, she is prompted for both her "pan-disease" (pan-protocol) global characteristics (e.g. demographic and historical phenotypic data such as age, performance status, smoking status, pregnancy status, etc) and her diagnosis (breast cancer). Based on her diagnosis, the CRFQ prompts her via a series of disease-specific questions to collect a variety of "disease-specific" phenotypic and genotypic data such as stage, HER2-Neu status, ER/PR status, prior chemotherapy, prior hormone therapy, etc. Based upon her responses, the CRFQ accesses a publicly-available Protocol Repository which contains the Inclusion/Exclusion (I/E) criteria for a number of currently open protocols for her disease. The CRFQ returns to the woman a list of all protocols for which she might be eligible based on an initial comparison of her specific data with the protocol's I/E criteria. The CRFQ interface also offers her an additional option to further restrict the candidate protocols on the list based on geographic proximity and/or a preferred radius of travel-to-treatment miles using a comparison of her zip code with the protocol's associated treatment site zip code(s). The CRFQ also enables her to restrict the list of candidate protocols for her to investigate further by allowing her to filter the list of protocols based on the type of intervention provided by the protocol (e.g. primary treatment, adjunct treatment, vaccine, etc.). The final list of candidate protocols contains a protocol-specific URL which, when traversed, provides the protocol-specific, full-study synopsis as posted on www.clinicaltrials.gov . The woman prints the final list to take for discussion with her physician.



3.2 Scenario #2: Protocols searching for Potential Subjects

A Trial Sponsor has developed a new intervention for Type I diabetes and has developed a clinical trial protocol to test this new intervention. A repository containing the Electronic Health Records (EHRs) for a number of patients is available to the Sponsor as a possible source of subjects for the protocol. The Trial Sponsor (or the Sponsor's designated Agent for querying a specific EHR repository) invokes the CRFQ service and enters in the "pan-protocol" (non-disease-specific) Inclusion/Exclusion (I/E) criteria required of each subject by the protocol (e.g. demographic or general phenotypic data such as gender requirements, age range, allowable smoking or alcohol use history, allowed/disallowed concurrent diagnoses, etc.), as well as the protocol's disease-specific diagnostic criteria (e.g. established diagnosis and category of Type I Diabetes). If appropriate, the CRFQ then prompts the Sponsor/Agent for a more detailed set of disease-specific I/E criteria (e.g. minimum time since onset, allowable past treatment history, required organ status, range of allowable insulin therapy etc.). The CRFQ then accesses the EHR repository and identifies all patients whose individual data satisfy the constraints specified by the protocol's I/E criteria. At this point, there are two branches in the Storyboard, depending on whether the Sponsor/Agent has established a 'Trusted' or 'Non-Trusted' relationship with the owner of the EHR repository. The status of the Client/EHR repository relationship is determined by the Security Service and passed to the CRFQ at the time of service invocation*:*

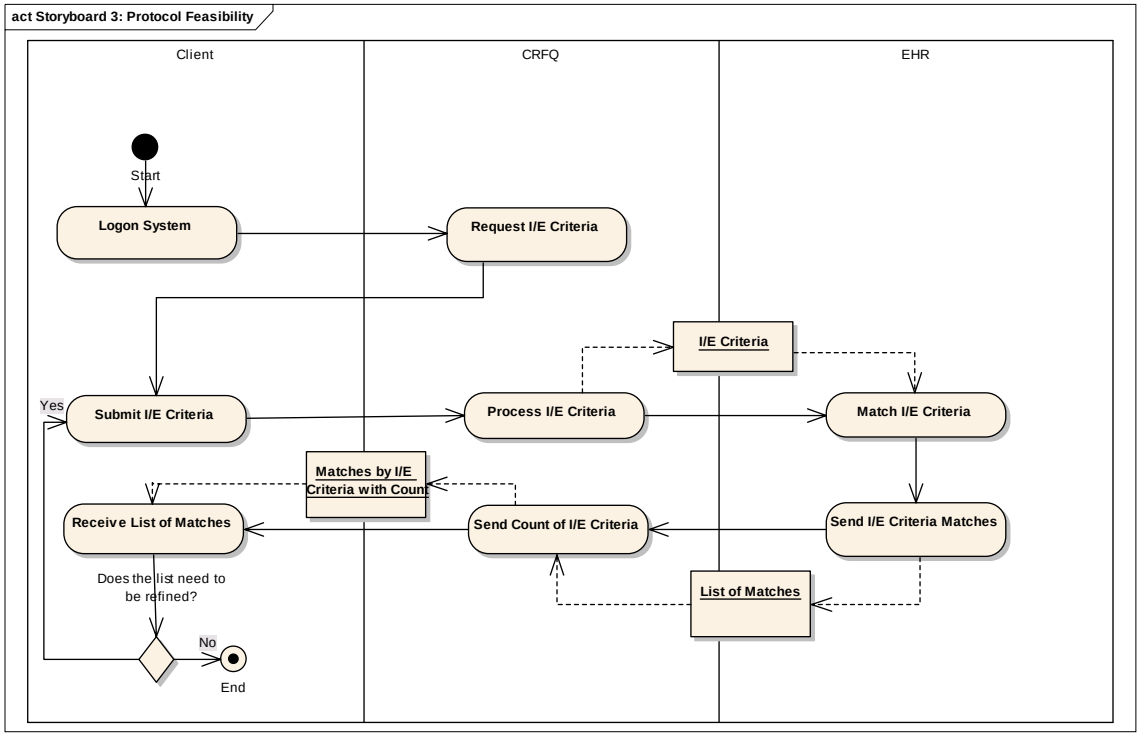


3.2 Scenario #3: *Sponsors evaluating I/E Criteria Efficacy*

a) A drug company is conducting a protocol of a new intervention in pediatric hypertension which requires 400 children as subjects. One of the protocols' I/E criterion is based on ranges of allowable lab values. After one year, the company has only recruited 50 children, half of the original estimated recruitment rate. The company (or their designated agent(s)) invokes the CRFQ service against a number of EHR repositories to which they have access to determine whether changing/broadening/relaxing the laboratory I/E criteria would allow the recruitment of more subjects otherwise qualified for the protocol.

b) A protocol is being developed and the basic characteristics of the patient population to be studied are known. There is, however, some uncertainty about how restrictive some of the protocol's I/E criteria can be relative to the identified potential subject pools that can be accessed by the protocol. The protocol developer wants to run several versions of the protocol's various I/E criteria against available EHR repository (ies) to see how many subjects would meet the different I/E criteria.

c) An Investigator has been approached about participating in an upcoming trial, and has been asked by the Trial Sponsor to estimate how many subjects the Investigator could recruit. The Investigator would like to participate in the study, but suspects that some of the I/E criteria are too restrictive. The Investigator needs to have quantitative data to back a request that the Sponsor revise I/E criteria. The Investigator invokes the CRFQ service against his/her local EHR repository using various modifications of the proposed I/E criteria to identify a set of I/E criteria which the Investigator believes are both scientifically reasonable and practically obtainable from the EHR repository.



733

734

735

736

4 Assumptions and Dependencies

Assumptions

The deployment context for the CFRQ service is based on the following assumptions:

- For *both interfaces, free-text data is not accepted unless the CRFQ client has the ability to parse the free-text into semantically non-ambiguous structures as required by the various CRFQ operations.*
- Furthermore, it is assumed that Inclusion/Exclusion criteria can be stated in semantically non-ambiguous grammars suitable for supporting automatic comparison of data elements (e.g. coded data from known coding systems, standardized representations of non-encoded data using HL7 V3 data type specifications, etc.) to enable automated processing of the criteria's semantics as described in the CRFQ service's Operations' specifications (Section 5).
- For the *List Qualified Protocols* interface (see Section 7 for system interaction diagrams), there are no known security, anonymization, or consent issues, i.e. *available protocols and their associated metadata will be stored in a publicly accessible repository to which the CRFQ service will have access.*
- For the *List Qualified Patients* interface, (see Section 7 for system interaction diagrams), there are several assumptions relating to the existence of data of sufficient semantic robustness including:
  - Appropriate security (authentication/authorization) service(s) exist;
  - Patient data with semantic robustness equivalent to that expressed by the I/E criteria;
- Ability to express 'patient preference data' e.g. desired treatment location, disease focus and intervention type desired (e.g. *breast cancer – chemotherapy*) in computationally non-ambiguous terms;
- Existence of manual, semi-automated, or automated consent management (aka 'trusted broker') service(s);
- The ability of the Security Service(s) to determine the 'degree of trust' of a CRFQ client and of the CRFQ service to modify its functionality accordingly. In particular, the CRFQ service depends on the Security Service(s) to verify a given client's trusted/non-trusted status and to pass this information to the CRFQ. The status of a CRFQ client determines the data returned to the client: *"if trusted and consent present, return list of potential patients, otherwise refer list to an "honest broker to pursue consent*; *if non-trusted, return count of potential patients.* The status of trusted/non-trusted also determines whether the CRFQ will require the cooperation of an anonymization/pseudonymization service as shown in the business activity diagrams (see Section 3).

775 • Existence of a protocol repository/repositories which include the ability to register
776 protocols and their associated PIs, etc.

777 Other assumptions on the nature of both the deployment and business context are
778 discussed in Section 2.

779

5 Functional Specification for each CRFQ Interface (operations supported by CRFQ interfaces)

The following operations support the business storyboards described in Section 3 (above). It should be noted that these operations fall into two implicit categories: mainline and supporting. The mainline operations are List Qualified Protocols and List Qualified Patients, while all other operations are supporting. The distinguishing characteristic between the two types of operations is that the mainline operations will always be required to fulfill the business cases while the supporting operations may or may not need to be called, depending on the capabilities of the consuming system. As this is a functional model, and since this delineation serves no business purpose, it is left implicit rather than providing a sub-categorization.

Additionally, note that CRFQ describes the RLUS service as a potential dependency when interacting with registries of either protocols or patients. Because of this dependency, CRFQ can be focused on solving the business problems associated with functional queries for clinical research, rather than the administration of repositories and registries.

5.1 List Qualified Protocols

(Note: (M) Mandatory and (O) Optional)

Description (M)	Based on patient-specific data supplied by the CRFQ client, the interface enables the identification of any protocols a particular patient qualifies for based on the each protocol's I/E Criteria ('protocol metadata'). Can optionally account for Patient Preferences and I/E Criteria-specific Weighting or Scoring metadata.
Precondition (M)	▪ Computable protocol metadata – available by reference or by value -- defining a protocol's I/E criteria ▪ Patient-specific data – available by reference or by value -- describing a patient's demographic, phenotypic, and genotypic profile in terms equivalent to the semantics of the protocol metadata
Inputs (M)	▪ (M) Disease-independent data for single patient ▪ (M) Disease-specific data for single patient ▪ (M) Protocol or safety-even Inclusion criteria ▪ (M) Protocol Exclusion or safety-event criteria (NOTE: It may be the case that only Inclusion <<or>> Exclusion data are available. One of the two data sets must, however, be present) ▪ (O) Criteria-specific Preferences ▪ (O) Scoring and Weighting Preferences Criteria
Outputs (M)	▪ (M) List of Qualified Protocols (may be factored by Patient Preferences and/or Weighting/Scoring metadata) ▪ (O) Element-by-element status values (may be factored by Patient or Scoring/Weighting Preferences): E.G. -- For *each protocol* in which *at least one* patient-specific data element met a protocol *Inclusion* criteria, list (for that protocol) all data elements in the input dataset and an associated status detailing: **element =/ Inclusion** (data value vs. inclusion value) **element = exclusion** (data value vs. exclusion value)
Post-Condition (O)	▪ A list of protocols for which a single patient may qualify based on input data/metadata is available to the service consumer

Exception Conditions (M)	▪ Patient-specific, disease-independent data not parsable/computable for comparison with I/E Criteria ▪ Patient-specific, disease-specific data not parsable/computable for comparison with I/E Criteria ▪ Protocol metadata (I/E Criteria) not parsable/computable for automated comparison to Patient-specific data ▪ Patient Preferences not parsable/computable ▪ Scoring and Weighting Preferences not parsable/computable

Aspects left to RFP Submitters (M)	▪ *Expression and/or scope of:*
	  ▪ Protocol metadata
	  ▪ Disease-independent data
	  ▪ Disease-specific data
	  ▪ Patient Preference metadata
	  ▪ Scoring and Weighting metadata

Relationships to levels of Conformance (or other patterns) (O)	Query for Qualified Protocols Functional Profile
Notes (O)	*Note that the use of I/E or safety-event Criteria may be implicit or explicit in the service interface. Essentially, there may be a single interface that is disease-/event-agnostic, or a single interface per disease/event, or some combination. This aspect should be specifically addressed by implementers.*

798

799

5.2 List Qualified Patients

Description (M)	Based on protocol-specific metadata describing a protocol's I/E Criteria -- or safety event-of-interest metadata describing a potential 'signal' condition – either set of metadata supplied by the CRFQ client, the interface enables the identification of any patients whose individual data satisfies the selection criteria. Can optionally account for Patient Selection Preferences and I/E Criteria-/safety-event-specific Weighting or Scoring metadata.
Precondition (M)	▪ Computable protocol metadata – available by reference or by value -- defining a protocol's I/E criteria ▪ Patient-specific data – available by reference or by value -- describing a patient's demographic, phenotypic, and genotypic profile in terms equivalent to the semantics of the protocol metadata
Inputs (M)	▪ (M) Disease-independent data for individual patients ▪ (M) Disease-specific data for individual patients ▪ (M) Protocol or safety-event Inclusion criteria ▪ (M) Protocol Exclusion or safety-event criteria (NOTE: It may be the case that only Inclusion <<or>> Exclusion data are available. One of the two data sets must, however, be present) ▪ (O) Criteria-specific Preferences ▪ (O) Scoring and Weighting Preferences Criteria
Outputs (M)	▪ (M) List of Qualified Patients (may be factored by Patient Preferences and/or Weighting/Scoring metadata) ▪ (O) Element-by-element status values (may be factored by Patient or Scoring/Weighting Preferences): E.G. -- For *each protocol* in which *at least one* patient-specific data element met a protocol *Inclusion* criteria, list (for that protocol) all data elements in the input dataset and an associated status detailing: **element =/ Inclusion** (data value vs. inclusion value) **element = exclusion** (data value vs. exclusion value)
Post-Condition (O)	▪ A list of patients who fit the I/E or safety-event Criteria, modulated by Scoring or Weighting metadata if available is made available to the service consumer

Exception Conditions (M)	▪ Patient-specific, disease-independent data not parsable/computable for comparison with I/E Criteria ▪ Patient-specific, disease-specific data not parsable/computable for comparison with I/E Criteria ▪ Protocol metadata (I/E Criteria) not parsable/computable for automated comparison to Patient-specific data ▪ Patient Preferences not parsable/computable ▪ Scoring and Weighting Preferences not parsable/computable
Aspects left to RFP Submitters (M)	▪ *Expression and/or scope of:* ▪ Protocol metadata ▪ Disease-independent data ▪ Disease-specific data ▪ Patient Preference metadata ▪ Scoring and Weighting metadata
Relationships to levels of Conformance (or other patterns) (O)	Query for Qualified Patients (Protocol) Functional Profile Query for Protocol Efficacy Functional Profile Query for Qualified Patients (Safety Event) Functional Profile
Notes (O)	*Note that the use of I/E or safety-event Criteria may be implicit or explicit in the service interface. Essentially, there may be a single interface that is disease-/event-agnostic, or a single interface per disease/event, or some combination. This aspect should be specifically addressed by implementers.* *The*

801

802 **5.3 List Inclusion Criteria**

Description (M)	The CRFQ Service is self-descriptive and can therefore can list the explicit Inclusion Criteria resident in the protocol repository (ies) which host the CRFQ Service and against which patient-specific data will be compared.
Precondition (M)	▪ Inclusion Criteria are Expressed in a parsable/computable format
Inputs (M)	▪ (O) Filter criteria for Inclusion Criteria
Outputs (M)	▪ List of Inclusion Criteria
Post-Condition (O)	▪ A List of Inclusion Criteria is returned to the service consumer
Exception Conditions (M)	▪ N/A
Aspects left to RFP Submitters (M)	▪ The parsable/computable expression of the Inclusion criteria
Relationships to levels of Conformance (or other patterns) (O)	Query for Qualified Protocol Functional Profile Query for Qualified Patients (Safety Event) Functional Profile
Notes (O)	*Different implementations may support different representational or semantic partitions/subsets of the universe of all possible Inclusion Criteria.* *Because a service specification can make very few assumptions about the nature of the consuming system, this operation exists to allow someone inventory the questions (5.6,5.7) to elicit the data (5.3, 5.4). So it is a different interface. These interfaces support the use of CRFQ in many different deployment contexts without specifying a particular one. For example, it may be that the Inclusion Criteria is an explicit parameter of a CRFQ operation, or it may be implicit (such as when standing up a 1 instance per set of I/E criteria).*

803

804

5.4 List Exclusion Criteria

Description (M)	The CRFQ Service is self-descriptive and can therefore list the explicit Exclusion Criteria resident in the protocol repository(ies) which host the CRFQ Service and against which patient-specific data will be compared.
Precondition (M)	▪ Exclusion Criteria are Expressed in a parsable/computable format
Inputs (M)	▪ (O) Filter criteria for Exclusion Criteria
Outputs (M)	▪ List of Exclusion Criteria
Post-Condition (O)	▪ A List of Exclusion Criteria is returned to the service consumer
Exception Conditions (M)	▪ This operation is not available in the case of the deployment of the CRFQ Service in a safety-event-monitoring context because of the absence of appropriate metadata.
Aspects left to RFP Submitters (M)	▪ The parsable/computable expression of the Exclusion criteria
Relationships to levels of Conformance (or other patterns) (O)	Query for Qualified Protocol Functional Profile Query for Qualified Patients (Safety Event) Functional Profile
Notes (O)	*Different implementations may support different representational or semantic partitions/subsets of the universe of all possible Exclusion Criteria*

807 **5.5 List Pan-Protocol Metadata**

Description (M)	The CRFQ Service is self-descriptive and can therefore list the explicit pan-protocol metadata resident in the protocol repository (ies) which host the CRFQ Service and against which patient-specific data will be compared.
Precondition (M)	▪ Pan-Protocol metadata are expressed in a parsable/computable format
Inputs (M)	▪ None
Outputs (M)	▪ List of Pan-Protocol metadata
Post-Condition (O)	▪ A List of Pan-Protocol metadata for the protocol repository (or repositories) hosting the CRFQ Service is returned to the service consumer
Exception Conditions (M)	▪ This operation is not available in the case of the deployment of the CRFQ Service in a safety-event-monitoring context because of the absence of appropriate metadata.
Aspects left to RFP Submitters (M)	▪ The parsable/computable expression of the Pan-Protocol metadata criteria
Relationships to levels of Conformance (or other patterns) (O)	Query for Qualified Protocol Functional Profile
Notes (O)	*Different implementations may support different representational or semantic partitions/subsets of the universe of all possible Exclusion Criteria*

808

809 **5.6 List Disease-Independent Questions**

Description (M)	The CRFQ Service is self-descriptive and can therefore list the specific disease-independent (pan-protocol) questions used to gather data for comparison against protocol metadata
Precondition (M)	▪ Disease-independent metadata is available – by value or reference – for each CRFQ instance
Inputs (M)	▪ None
Outputs (M)	▪ List of disease-independent (and pan-protocol) questions that can be used to gather patient-specific data for comparison to protocol-metadata
Post-Condition (O)	▪ The disease-independent questions are available to the CRFQ service consumer
Exception Conditions (M)	▪ The disease-independent metadata are not available in the protocol repository(ies) hosting the CRFQ Service
Aspects left to RFP Submitters (M)	▪ The parsable/computable expression of the disease-independent metadata
Relationships to levels of Conformance (or other patterns) (O)	Query for Qualified Protocol Functional Profile
Notes (O)	*Because a service specification can make very few assumptions about the nature of the consuming system, this operation exists to allow someone inventory the questions (5.6, 5.7) to elicit the data (5.3, 5.4). So it is a different operation. These interfaces support the use of CRFQ in many different deployment contexts without specifying a particular one.*

810

5.7 List Disease-Specific Questions

Description (M)	The CRFQ Service is self-descriptive and can therefore list the disease-specific questions used to gather data for comparison against protocol metadata (listed by disease)
Precondition (M)	▪ Disease-specific metadata is available – by value or reference – for each CRFQ instance
Inputs (M)	▪ (O) Filter Criteria (per disease)
Outputs (M)	▪ List of disease-specific questions that can be used to gather patient-specific data for comparison to protocol-metadata
Post-Condition (O)	▪ The disease-specific questions are available to the CRFQ service consumer
Exception Conditions (M)	▪ The disease-specific metadata are not available in the protocol repository(ies) hosting the CRFQ Service
Aspects left to RFP Submitters (M)	▪ The parsable/computable expression of the disease-specific metadata
Relationships to levels of Conformance (or other patterns) (O)	Query for Qualified Protocol Functional Profile
Notes (O)	*Because a service specification can make very few assumptions about the nature of the consuming system, this operation exists to allow someone inventory the questions (5.6,5.7) to elicit the data (5.3, 5.4). So it is a different interface. These interfaces support the use of CRFQ in many different deployment contexts without specifying a particular one.*

6 Profiles

6.1 Introduction

A profile is a named set of cohesive capabilities. A profile enables a service to be used at different levels and allows implementers to provide different levels of capabilities in differing contexts. *Service-to-service interoperability will be judged at the profile level and not the service level. Note that through the use of profiles, there are no "optional" interfaces. Conditions that might otherwise merit this optionality should be addressed via a dedicated profile. Following are descriptions of Functional, Semantic, and Conformance Profiles for the CRFQ Service.*

6.2 Functional Profiles

A Functional Profile defines the specific operations of an interface that are used in a particular business context. In the case of the CRFQ service, the *List Qualified Protocols* business context (Scenario #1) utilizes all of the operations available in its interface. On the other hand, the *List Qualified Patients* interface is applied in three business contexts – Scenarios #2, #3, and #4, each application (potentially) defining a separate business context and therefore a separate Functional Profile. The Functional Profiles for the CRFQ are enumerated in the following table:

Functional Profile	Member Operations	Operation Profile	Notes
Query for Qualified Protocols for a specific patient based on patient-specific data	List Qualified Protocols	If protocols are expressed as parameters from the service provider, multiple protocols MUST be expressed as ordered collections with order based on number of Inclusion or Exclusion Criteria satisfied	Assumption is that the semantics and grammar of patient-specific data are compatible with the semantics and grammar of protocol metadata, thereby enabling automated comparison of the data sets

Functional Profile	Member Operations	Operation Profile	Notes
	List Inclusion Criteria (or safety-event-of-interest inclusion criteria)	If protocols are considered to be provided from the service provider, then the "List Inclusion Criteria" operation MUST provide all inclusion criteria for all protocols based on expressed patient preferences such as disease or treatment focus, location, etc.	
	List Exclusion Criteria (or safety-event-of-interest exclusion criteria)	If protocols are considered to be provided from the service provider, then the "List Exclusion Criteria" operation MUST provide all inclusion criteria for all protocols based on expressed patient preferences such as disease or treatment focus, location, etc.	
	List pan-protocol metadata	Same as functional description	
	List Disease-independent Questions	Same as functional description	
	List Disease-specific Questions	Same as functional description	

Functional Profile	Member Operations	Operation Profile	Notes
Query for Qualified Patients per protocol-specific Inclusion / Exclusion Criteria Set	List Qualified Patients, List Inclusion Criteria, List Exclusion Criteria	If patient population data is made explicit as an OUT parameter of the interface, it MUST be an ordered collection. If patient-by-patient status values are made available as an output and scoring and weighting preferences are used as an input, then the status values MUST be tied to the scoring and weighting preferences	
Query for patients with data matching a Safety Event signal	List Qualified Patients	If patient population data is made explicit as a parameter of the interface, it MUST be an ordered collection. If patient-by-patient status values are made available as an output and scoring and weighting preferences are used as an input, then the status values MUST be tied to the scoring and weighting preferences	

Functional Profile	Member Operations	Operation Profile	Notes
	List Inclusion Criteria (or safety-event-of-interest inclusion criteria)	Inclusion and Exclusion criteria would be related to events rather than strictly to protocols, though the format and the computational algorithm would be the same as in List Qualified Patients. If I/E criteria are provided by the service provider, these MUST be made available	
	List Exclusion Criteria (or safety-event-of-interest exclusion criteria)		
Query for Efficacy of Inclusion / Exclusion Criteria	List Qualified Patients	The input parameter "Scoring and Weighting Preferences on Inclusion and Exclusion Criteria" MUST be mandatory The output parameter "Patient-by-patient status values" MUST be mandatory	

6.3 Semantic Profiles

A Semantic Profile identifies a named (and robustly defined) set of data/information descriptions (e.g. semantic signifiers) that are supported by one or more operations. As described in Section 2, the CRFQ Service will support two interfaces:

- List Qualified Protocols
- List Qualified Patients

The following is an overview of the expected semantic profiles for these two interfaces from the perspectives of business context, localization, information models, partner-to-partner interoperability contexts, and anticipated product packaging and offerings.

6.3.1 Business Contexts

6.3.1.1 List Qualified Protocols

As discussed in Section 2, the primary business context for this interface (see Scenario #1) is consumer health access (i.e. patients or their providers searching for appropriate protocols for evaluation for participation). It is expected that patient-specific data will be expressed as a combination of BRIDG concepts (e.g. PerformedObservation, PerformedProcedure, PerformedMedicationAdministration, PerformedAssessment, etc.) bound, where applicable, to appropriate concept-based terminologies, e.g. LOINC, SNOMED CT, etc. (See the previous discussion on the specific, documented and mapped relationship between BRIDG Model semantics and RIM semantics.)

6.3.1.2 List Qualified Patients

As discussed in Section 2, there are two primary business contexts: clinical trial sponsors, investigators, and/or others interested in recruiting potential subjects into a clinical trial and/or refining a trial's proposed I/E Criteria; or persons/organizations interested in screening patient populations for the presence or absence of one or more safety event-of-interest ("signals.").

6.3.2 Localization

Both CRFQ Service interfaces can support localization with the recognition that locally encoded protocol metadata and/or patient-specific data degrades the interoperability of the underlying repositories. Likewise, failure to bind local repositories to V3 data types undermines the computable semantic interoperability of the data when it is transferred between systems.

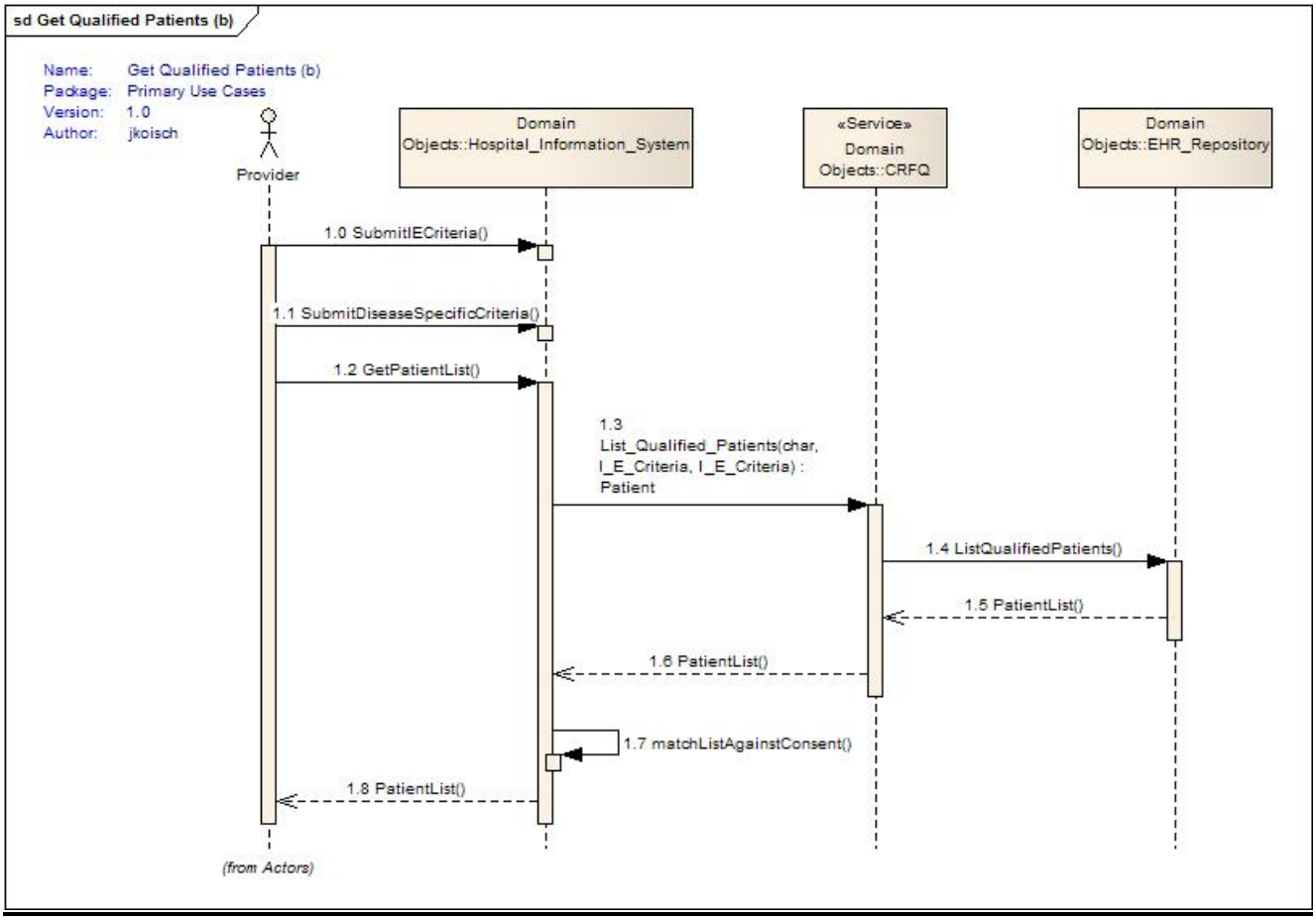
6.3.3 Information Models

Both interfaces of the CRFQ Service are expected to utilize the same information models or subsets thereof. In particular:

- all patient data, pan-protocol and protocol-specific metadata, and safety-event metadata will be bound to the appropriate HL7 Version 3 Abstract Data Type;
- all patient-specific data will be represented using a combination of BRIDG information concepts bound (when appropriate) to relevant terminologies (e.g. LOINC, SNOMED, etc.)

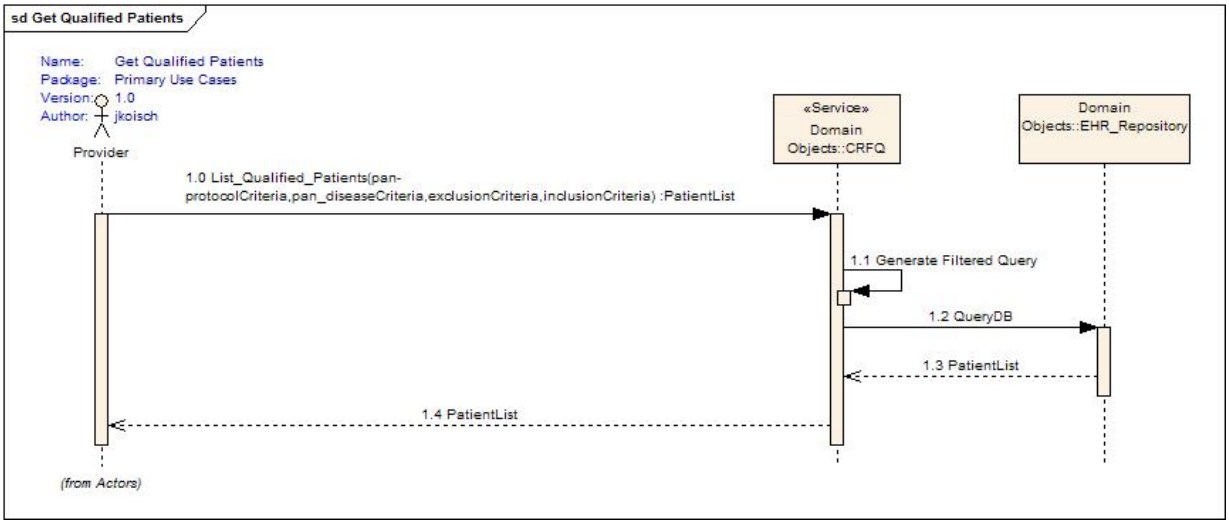
7 User Scenario Interaction Details

7.1 Scenario #1: Patients/Providers searching for Protocols



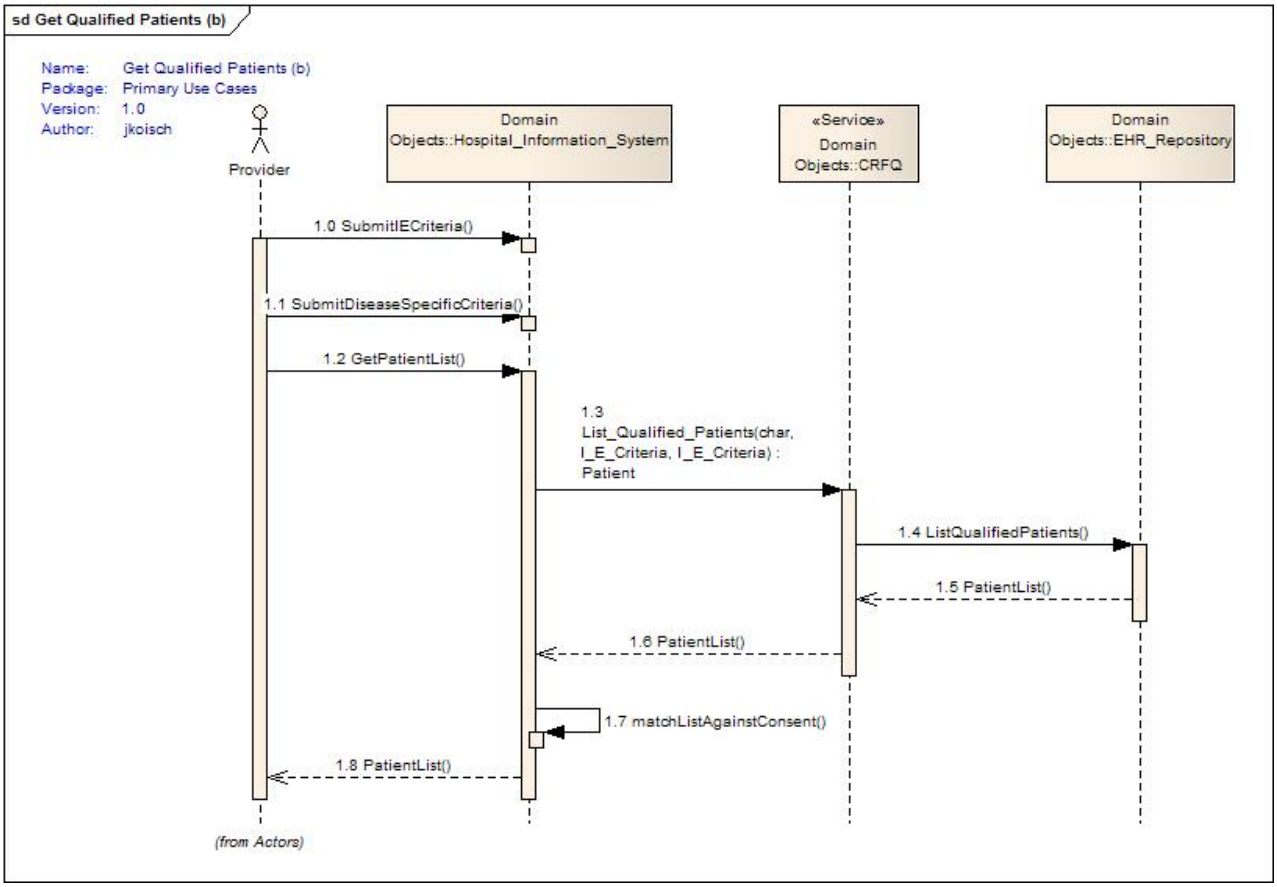
***List Qualified Protocols* using Patient-centric data and preferences**

7.2 Scenario #2: Protocols searching for Potential Subjects



List Qualified Patients **by inputting Inclusion and Exclusion Criteria, along with other general information**

7.3 Scenario #3: Sponsors evaluating I/E Criteria Efficacy



Evaluate Protocol Efficacy by measuring the effectiveness of Inclusion and Exclusion Criteria against patient populations using the *Get Qualified Patients* Operation

8 Relationship to Information Content: SFM Principles

The following principles shall be followed for specifying the information model to be used by the services being specified in this Service Functional Model:

1. SFMs shall provide a conformance profile supporting HL7 or other appropriate standards-based content where relevant
2. The SSF does not preclude the use of non-HL7 content
3. SFMs will reuse to the maximum extent possible the content models as defined in other standards (for example, HL7 Refined Message Information Models (RMIMs))
4. Information content representations shall be represented in platform-agnostic formalisms (e.g., UML)
5. SFMs may identify content at varying levels of granularity, depending upon the functions being specified.

6. Conformance Profiles may be balloted or adopted after the release of the initial SFM to address specialized business needs. (realm-specific profiles, domain-specific profiles, etc.)

9 Recommendations for Technical RFP Issuance

As referenced in Section 1, this SFM is intended to provide the functional model as the foundation for an RFP to be issued by the Object Management Group (OMG). The following issues and considerations are considered important by the HL7 community (RCRIM in particular) as responders evaluate this SFM for content and, ultimately, issue a technical specification based on this SFM.

While the HL7 community at large is invited to participate in the ongoing RFP process as subject matter advisors, the following open issues are left to submitters to the RFP.

9.1 Standardizing the Parameters using Semantic Signifiers

RFP Responders should specifically address the use of semantic signifiers and their explicit expression for the following parameter types (as exemplified in the ASPIRE data set):

1. Demographic Data
2. Patient Disease Historical Data
3. Disease MetaData
4. Disease Data
5. Protocol Listings
6. Protocol metadata
7. Inclusion/Exclusion Criteria
8. Patient Preference Data
9. Scoring and Weighting metadata

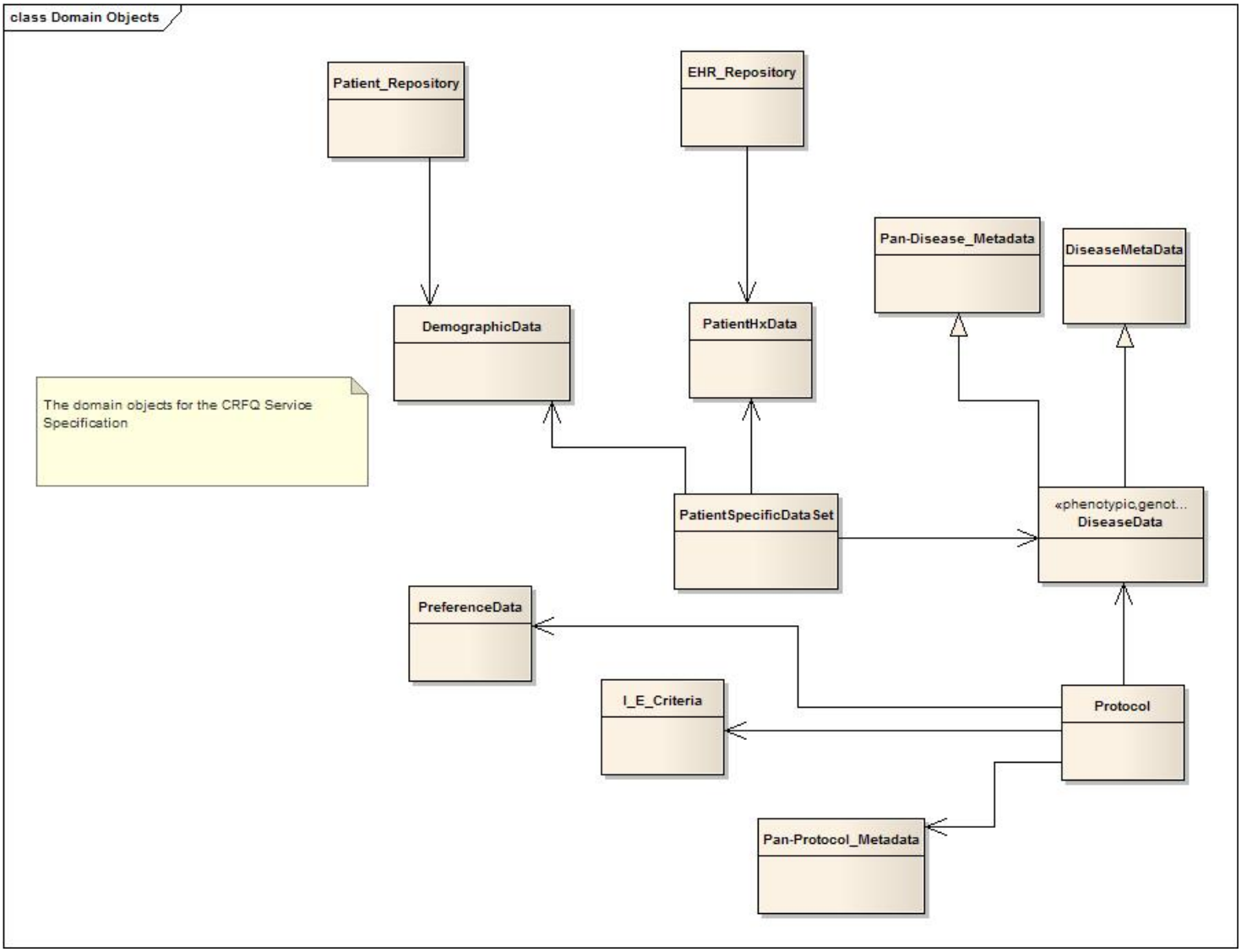


Figure 2: Relation of Domain Components (Informational)

9.2 Protocol/Disease Metadata Registries

RFP Responders should address the dependency of the CRFQ service on protocol and disease metadata, and the necessity of such metadata being available at run time to both the service implementation of the service and optionally to the client. Specific areas to address include performance, service visibility, semantic clarity, and whether metadata is available by reference or by value.

9.3 Metadata and Semantic Signifiers

Particular attention should be paid to the representation of metadata within the CRFQ context in light of the work done by RFP submitters on the RLUS project, particularly with respect to semantic signifiers. It is possible and even likely that metadata and informational components for CRFQ can be expressed at runtime using the semantic signifier notion.

9.4 Service Interfaces and Protocol Inclusion / Exclusion Criteria

RFP Responders should address the use of I/E Criteria as a parameter on various service as opposed to making the criteria implicit in the interface. In other words, the topic of interface deployment **per** disease or **per** protocol should be addressed explicitly.

Submitters to the RFP should be especially mindful of the issues of manageability, maintainability, and other operational concepts when describing the recommended way that CRFQ services could or should be provisioned.

9.5 Other HSSP Services

The RFP Response should specifically address the relationship between CRFQ and other HSSP services, especially including EIS, RLUS, and DSS. For example, it may be appropriate to express the List Qualified Protocol operation could be implemented as a DSS profile.

10 Appendix A - Relevant Standards

The three relevant standards are the RLUS Service, the BRIDG Model (which is technically only a *de facto* standard) and HL7 Version 3 Abstract Data Type Specification.

10.1 Retrieve, Locate, and Update Service

The Retrieve, Locate, and Update Service (RLUS) provides functionally consistent capability to multiple information models. While seemingly simplistic, RLUS provides a mechanism for categorizing and accessing information in a variety of deployment contexts.

For CRFQ, RLUS may provide a consistent means of accessing both protocol and patient data as composable elements "behind" the CRFQ service interfaces. Whether this is appropriate for a particular implementation would be determined locally, but for the purposes of envisioning CRFQ, this provides a consistent and reliable means of expressing information.

For more information about RLUS, see the Draft Standard for Trial Use available through HL7 at this location (http://www.hl7.org/dstucomments/index.cfm).

10.2 Decision Support Service

The Decision Support Service (DSS) receives patient data as the input and returns patient-specific conclusions as the output. DSS provides access to machine-executable medical knowledge though a service interface and initial was envisioned to facilitate the implementation and maintenance of Clinical Decision Support (CDS) capabilities within clinical applications. CRFQ will provide a foundational component for other services as both service consumer and provider, including DSS.

For more information about DSS, see the Draft Standard for Trial Use available through HL7 at this location (http://www.hl7.org/dstucomments/index.cfm).

10.3 The BRIDG Model

The BRIDG Model (http://www.bridgmodel.org), is a Domain Analysis Model (DAM) whose scope is "protocol-driven research involving human, animal, or device subjects and all associated regulatory artifacts".

10.4 HL7 Version 3 Abstract Data Type Specification

The HL7 Version 3 Abstract Data Type Specification (http://www.hl7.org) is an American National Standards Institute (ANSI) **s**tandard which defines the semantics of a collection of complex data types (e.g. Physical Quantity, Coded Description, General Timing Specification, etc.) which HL7 has found to be essential if machines are going to exchange data at a computable semantically interoperable level.

11 Appendix B – Glossary of Terms (in the context of CRFQs)

Term	Definition
Criterion (Criteria)	One (or more) statements/facts against which another statement/fact can be compared with the result being a True or False condition. The result may – but need not be – also quantitatively measurable. For example, the criterion "IsFemale" will be either True or False, whereas the criterion "Serum Na > 140 mEq/ml" may be "False by 9 mEq/ml," i.e. the serum Na is 149 meE/ml
Criteria, Exclusion	The set of criterion used to collectively determine whether a candidate subject should be excluded from consideration in a specific protocol's cohort. Thus, the Exclusion Criteria are considered metadata of the protocol. For example, "Protocol will not accept Females over age 50 or those with positive mammograms within the last 6 months."
Criteria, Inclusion	The set of criterion used to collectively determine whether a candidate subject should be excluded from consideration for inclusion in a specific protocol's cohort. Thus, the Inclusion Criteria are considered metadata of the protocol. For example, "Females between 35-50 with negative mammograms within the last 6 months."
Criteria, Scoring and Weighting	Data collected by the CRFQ which can be applied against specific Inclusion or Exclusion criteria when screening a particular patient data set against these criteria. Scoring and Weighting criteria are used by protocol administrations, clinicians and patients to determine the degree of influence that a particular criteria has on patient inclusion or exclusion, and/or as criteria for exploring 'closeness of fit' between a patient/potential subject and a given protocol.
Disease-Specific Data	Patient-specific data bound to one or more data elements that have been defined to be potential Inclusion or Exclusion criteria for one or more protocols with a disease-specific focus and are therefore used to gather patient-specific responses for comparison to a protocol's metadata. For example, disease-specific data for protocols with a focus on cardiac disease might include history of previous myocardial infarction, angioplasty, or CABB procedures; current cardiac medications; and cardiac hemodynamics.
Disease-Independent data	Patient-specific data bound to one or more data elements that have been defined to be potential Inclusion or Exclusion criteria for all protocols regardless of disease-specific focus and are therefore used to gather patient-specific responses for comparison to each protocol's metadata. For example, age, gender, ethnicity, etc.
Disease-Specific Questions	Questions with a disease-specific focus designed to elicit disease-specific data from patients, e.g. h/o cardiac illness, cardiac medications, etc.
Disease-Independent Questions	Questions with no disease-specific focus designed to elicit general data from patients, e.g. age, gender, ethnicity.

Term	Definition
Patient Preferences	One or more patient-specific data elements which can be used to restrict the results of a CRFQ List Potential Protocols service call, e.g. preferred protocol location, preferred protocol scope, etc.
BRIDG Model	A Domain Analysis Model (DAM) for the domain of "Protocol-driven research and its associated regulatory artifacts". The BRIDG model includes the data, organization, resources, rules, and processes involved in the formal assessment of the utility, impact, or other pharmacological, physiological, or psychological effects. The target of the model includes drugs, procedures, processes, or devices on a human, animal, or other biologic subject or substance.
ASPIRE Data Set	A collection of ~20 core coded eligibility proposed standard "pan-disease" data elements that cut across most protocols, particularly therapeutic studies, regardless of the disease entity under study (e.g. minimum and maximum allowable age, allowable gender(s), minimum performance status, etc.). Additionally, there are sets of "disease-specific" coded eligibility data elements that are shared among a specific disease entity (e.g. for breast cancer: stage of the cancer, estrogen and progesterone receptor status, HER2-Neu status, etc.)
Safety -related-event-of-interest (possible Adverse Event)	(CRFQ can be used to search patient populations for potential adverse events.) A safety-related-event (or reaction)-of-interest or is any adverse (or portentially adverse) change in health or "side-effect" that occurs in a person or animal. Is the consequence, (perhaps strongly indicated or only possibly related) of an investigative subject's use of a particular substance or product? For a person who participates in a clinical trial, a safety-related-event (or reaction)-of-interest is any adverse (or potentially adverse) change in health or "side-effect" that occurs in a while the patient is receiving the treatment (study medication, application of the study device, etc.) or within a pre-specified period of time after their treatment has been completed. Safety-related-events (or reactions)-of-interest in patients participating in clinical trials may be reported to the local Institutional Review Board (IRB) and the study sponsor. Safety-related-events (or reactions)-of-interest may subsequently be categorized as "serious" (for example death, illness requiring hospitalization, events deemed life-threatening, or involving cancer or fetal exposure) and may need to be reported to the regulatory authorities immediately, whereas minor events/reactions may simply be documented in the annual summary which may be sent to the regulatory authority.
Protocol	A structured sequence of steps, activities, observations, etc. usually linked by rules, designed to answer a proposed hypothesis.

Term	Definition
HL7 V3 Data Type Specification	An ANSI-certified specifications defining the semantics of a number of Abstract Data Types (i.e. the semantics are specified in an implantation-independent manner). The data types specified in the ANSI standard vary from relatively 'simple' data types (e.g. name, address, etc.) to more 'complex' data types (e.g., Physical Quantity, General Timing Specification, etc.). Each attribute in the BRIDG Model is bound to a specific HL7 V3 Data Type as a basic underpinning for achieving computable semantic interoperability.

986