

標準健診フォーマット実装ガイドの開発

木村 映善^{*1}, ^{*2} 窪寺健, ^{*3} 長瀬嘉秀

^{*1} 愛媛大学 ^{*2} 日本医師会総合政策研究機構, ^{*3} テクノロジックアート

Development of Implementation Guide for general check-ups

^{*1}Eizen Kimura, ^{*2}Takeshi Kubodera, ^{*3} Yoshihide Nagase

^{*1} Dept. of Medical Informatics, Medical School of Ehime University

^{*2} Japan Medical Association Research, ^{*3} Technologic Arts

抄録: データヘルスに貢献する高品質な健診データを提供頂くためには、健診項目や単位を標準化し、かつデータ発生源に近い段階でバリデーションを実施することが求められる。現状では日本医学健康管理評価協議会が健診の項目名、結果の値、単位、正常範囲、結果値相関を一体的に管理する標準規格として「健診標準フォーマット」の策定をすすめている。本研究ではこの健診標準フォーマットをベースに、これまでの標準規格に関する取り組みと比較して、健診データの高品質により寄与しうるFHIR実装ガイドを設計した。統制用語集とバリデーションロジックをデータベース化し、計算処理可能な実装ガイドを自動生成するスキームを設計することで、総合運用性を担保する規格の開発期間の短縮と高品質化に寄与可能であることが知見として得られた。

キーワード 標準健診フォーマット、FHIR、バリデーション、ターミノロジー

1. はじめに

法律で義務づけられている特定健診以外の諸々の健診結果については、JAHISの健康診断結果報告書規格はあるものの健診機関には費用・作業負担が大きく、健診システムベンダーにとっても標準化に対するインセンティブがないことで広く普及するに至っていない。2008年に健診事業者を構成員とする日本医学健康管理評価協議会が設立され、健診標準フォーマットの検討が開始された。データ提出者側の開発容易性・閲覧性を重視して表記を標準化し、CSVファイルを基本構造としている。将来における意味論的互換性の担保やバリデーションは規格の外側で検討する構図である。リアルワールドデータを用いた研究ではデータソースでのデータ品質担保の取り組みが重要であることは論をまたないが、データ発生源での対策は途上である。

現状では相互運用に関する標準規格の実装ガイド(以下、IG: Implementation Guide)は自然文章で記述されていることが主流である。またIG準拠を検証する検証プログラムの開発者はIG開発者と異なる第三者によって開発されることがある。IGの策定者とIG検証プログラム開発者、IGに準拠したデータ実装者の解釈のすり合わせには時間がかかり、検証プログラムの開発、データ

提出者への配布・データ提出者によるIGにもとづいた実装、検証プログラムを用いての検証と評価のサイクルが遅延する原因となる。そこで、健診標準フォーマットの策定活動と併走して、健診標準フォーマットを採用した事業者が次の段階においてデータ品質の向上やPHRに供するデータを開発するための検証手段として活用できる、計算処理可能なIGの開発を着想した。計算処理可能とは、実装ガイド上における規則の記述について自然文章の記述を極力排し、コンピュータによって規則に準拠していることを判定できるような評価式中心に記述することである。

2. 方法

健診標準フォーマットの情報モデルと標準コード表マスタ、標準健診項目とFHIR Resourceの各属性との対応と英訳、入力上限下限参考値、JCCLS 共用基準範囲、単位の標準化コード Unified Code for Unit of Measure (UCUM)の対応、QRコード(1)の情報を保有する情報モデルを設計し、RDBMS上に展開する。健診結果報告書を構成するリソースに関して、Profileを定義するドメイン固有言語であるFHIR Shorthand (fsh)を利用して定義した。FHIRのプロファイルでは、リソースの要素に関する制約を'Invariant'と称し、FhirPath式を用いて制約を記述する。多数の

Invariant がある Observation と Bundle において、Invariant を構成するパラメータを RDBMS から抽出し、Invariant 規則を記述する fsh スクリプトと FHIR Path 式を生成するプログラムを開発した。また、RDBMS から標準健診フォーマットの項目及びコード表に関する情報を抽出し、CodeSystem、ValueSet を生成する fsh スクリプトを自動生成するツールを開発した。自動生成された統制用語集、リソースのプロファイルを HL7 協会が配布している IG Publisher でコンパイルし、IG を生成した。

3. 結果

IG に実装された標準健診項目の CodeSystem は予約分も含めて 1500 語を収録、カテゴリカルデータの語彙を管理する CodeSystem は 84 種、総 2808 語の収録となった。各 Invariant に関して、カテゴリカルデータ記述に使う CodeSystem の制御が 398 種、データ型の規則は 397 種、入力上限下限参考値の規則は 106 種、UCUM Unit の規則が 477 種、結果値関連の規則が 24 種、JCCLS 共用基準範囲で実装された規則が 35 種となった。検査結果に関する Invariant 群は Observation、結果値関連と JCCLS 用基準範囲に関する Invariant は Bundle に実装された。健診結果を記述するために Bundle、Composition、Observation、Patient、Organization、Coverage の Profile が作成された。IG Publisher で生成されたバリデータツールを用いて、検査結果を記述している FHIR リソースに対してバリデーションを実行し、検証結果が HTML 形式で報告されることを確認した。

4. 考察

HL7 ver.2 は計算処理可能な規則を記述する手段は提供されていない。HL7 ver.3 は XML Schema、Schematron による記述が可能だが、特定 ValueSet へのバインディング記述や、スライシングを利用した適用プロファイルの制御、臨床結果関連チェックの記述は困難である。

結果値を特定のユースケース用に設計された

語彙集に収容されているものに限定するには、特定 ValueSet 由来の値に制約するバインディング制約が使われる。結果値関連チェックとあわせて、一定品質のデータを提出させるのに有効な手段であると思われる。今回は、複数プロファイルを組み合わせられる能力を利用していないが、NeXEHRs で策定している日本向けプロファイルや臨床検査のプロファイルが策定されれば、これらのプロファイルと組み合わせた IG を設計することで、医療機関由来の検査値と統合したライフコースデータを実現するような相互運用性に貢献することが期待される。また、FHIR は XML に加えて JSON、RDF をサポートしており、モバイルデバイスへの展開が期待される PHR やセマンティック Web 技術を利用した横断的な医療情報の活用への用途を拓いていくことが期待される。

5. 結語

本研究は計算処理可能な FHIR Profile を仕様リポジトリから自動生成する手法について考察、開発、検証した。本手法によって IG 開発者と IG 準拠検証プログラムの開発者を一体化し、開発プロセスの加速と様々な検証ロジックを実装することによるデータ品質向上の可能性を提示した。また、FHIR の Profile に関する能力は今後多様化、精緻化する検査結果を正確に収集するのに資することが確認された。「計算処理可能」とは、IG に関する規則を自然文章ではなくコンピュータが解釈して処理可能な様式で記述し、かつその記述から検証ロジックや検証をするツールを自動生成することが可能な状態であり、人手による解釈の曖昧性や実装時の人為的エラーを排除することが期待され、1000 項目以上に渡る巨大な IG 開発時には効果的であることが期待される。一方、由来情報の標準化、JLAC10 のハーモナイズにおいて課題が確認された。今後の健診標準フォーマットの IG の実装を進めていくにあたり、JLAC10 の規格にはない、LOINC であるところの Answer List について整備しつつ、JLAC10 とのマッピングについて検討を進めたい。翼, 阿部 達也. 健診標準二次元コード 特設ページ 2021 [Available from: <https://htech-lab.co.jp/normalization/>].

参考文献

1. 石見 拓, 島本 大也, 澤田 砂織, 高橋