

# 「知識の発見とデータ処理手順の可視化を支援する VRP」

紀ノ定保臣

岐阜大学大学院医学系研究科医療情報学分野

“ビッグデータの活用”が社会の重要なキーワードとなっている。また、医療分野では、日常の診療行為のみならず、医療技術の向上を目指した臨床研究や新しい医薬品開発のための臨床試験が活発に実施されている。患者を対象とした臨床研究等の場合、医療現場で発生する膨大な診療記録等のデータを活用することが一般的であり、データに対する信頼性（真正性）や保存・管理・分析手順に対する透明性が極めて重要となる。また、“医療の質保証”に対する関心が高まり、電子カルテシステムや診療部門システムのデータをデータ・ウェアハウス（DWH）化し、医療の質向上や臨床指標の作成・公開、病院マネジメント等に活用する事例が増加している。

一方、大規模に蓄積されたデータを効果的に分析・活用するためのデータマイニング手法に関する書籍は多くあるものの、実際の診療記録等データを用いて分析する、あるいは分析手法を開発できる人材は十分に育っていない。また、多くの医療機関が作成・ホームページ上に公開している臨床指標等については、その臨床指標の定義や作成手順、さらには臨床指標の算出に利用された原データ（データ項目名）が明確でないという課題が指摘されている。同様に、複数の医療機関で同一の臨床指標名称が用いられている場合、これらは比較可能であるかどうかも現状では不明瞭であるという課題がある。

本研究では、上記の課題を解決する目的で以下に示す3つのテーマに取り組むこととした。

(i) 医療従事者による実施・介入行為等データと診療の流れに沿って発生した患者毎の診療記録等データを DWH 化し、「医療の質保証」と「データの質保証」を支援するシステムとして再構築・運用する。

(ii) DWH に格納されたデータを対象に、医療の質向上や臨床指標の作成と公表に活用できる標準的な Reproducible Research 環境（臨床指標作成の手順と分析結果やグラフが常に再現できる環境、<http://cran.r-project.org/web/views/ReproducibleResearch.html> を参照）を構築する。

(iii) 標準化された DWH と Reproducible Research 環境を用いて、医療分野のデータサイエンティストを育成（計画）する。

研究開発対象となるシステムの運用概念図を図1に示す。DWH では個々の患者毎に別途定められる医療の質評価基準に基づいた質レポートが作成・蓄積・抽出され、分析対象症例群毎に医療の質についての統計的計算が実施され、組織としての医療の質に関する報告書を作成・公表するシステムとして完成を目指す。尚、大規模 DWH に対する検索を 30 秒以内に完了することができれば、研究者の思考を停止させることなく患者の臨床的・フローを対象とした「医療の質保証」に必要な臨床指標等を対話的・探索的に作成することが可能になり、システムとしての実用性が高まると期待される。

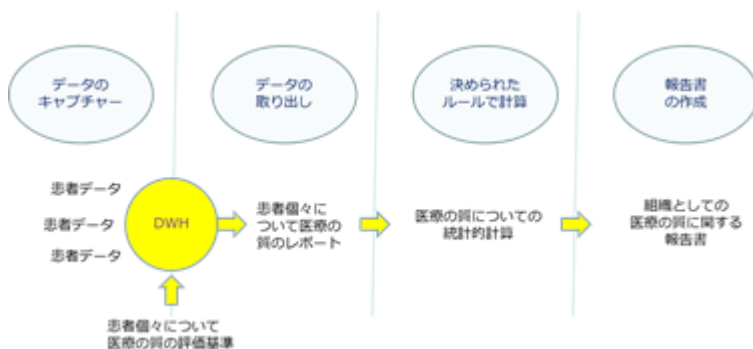


図1. Quality Report 作成手順