

CRDS-FY2010-GR-01

特定課題ベンチマーク報告書
トランスレーショナル・
ヘルスインフォマティクス・ベース

Benchmarking Report

Translational Health Informatics Base



エグゼクティブサマリー

優れたライフサイエンスの成果を活用し、新しい治療法や医薬品・医療機器等として社会に還元していくための「健康研究（Health Research）」（橋渡し研究・臨床研究）の必要性が謳われている。我が国においては、2008年に設立された健康研究推進会議（内閣府）が「健康研究推進戦略」を発表し、基礎研究により生命現象と疾病のメカニズムを解明し、それを診断や治療法に転換して医療を実践し、その医療効果を評価して新たな課題を設定し、それを再び基礎研究につなげていく、医学・医療における知の循環の確立を目指すべきであるという問題意識が示されている。

上記を踏まえ、独立行政法人科学技術振興機構 研究開発戦略センター（JST-CRDS）では、戦略プログラム『生命・医学・医療・健康をつなぐ情報を循環させる技術と基盤の構築と活用 ～トランスレーショナル・ヘルスインフォマティクス・ベースの展開～』を通して、医学・医療における「知の循環」の確立に必要な情報を循環させる技術と基盤（ここでは THIB：トランスレーショナル・ヘルスインフォマティクス・ベースと名づけた）の構築と活用を提案した。さらに、上記提案内容に類似した取り組みを先んじて行っている欧米の組織と比較し、より強力な研究開発推進体制の具体化の検討に資するために、国際ベンチマーク調査を実施した。本国際ベンチマークでは、①ライフサイエンス、ヘルスサイエンス双方における大容量データの収集、保管、抽出技術の現状②データ収集における提供者側へのインセンティブの付与戦略、③医療情報を基礎研究あるいは臨床現場に活用するための技術的制度的課題の克服、について、関係スタッフによる調査分析を行ったほか、当該分野の研究に詳しく、且つ海外の研究者コミュニティとの幅広いネットワークを持つ専門家に国際ベンチマーキング参加を依頼し、特に、米国における医療情報の管理、活用形態、ならびにドイツと英国におけるバイオインフォマティクスとヘルスサイエンスの融合研究プロジェクトと情報インフラ整備の現地調査を通じて我が国との比較を行った。その結果、以下のような結果が得られた。

<米国>

米国では、大規模な医療情報データベースを基盤とした産業が発達している。こうした産業の主要なプレイヤーとしては保険会社、IT企業の双方があり、これらが行政や医療機関、さらには研究機関や国民向けに情報を基盤としたサービスを提供している。わが国とは異なり、米国の医療制度では民間企業が医療保険財政にも大きく関わっているため、大規模医療情報を用いてより効率的、効果的な医療サービスを実現する、というニーズが産および官で共有されている印象が強い。あわせて、医療情報の扱いについても、個人情報保護の観点とその有意義な利活用の観点の両面から規制が整備されている。また、これら情報基盤の構築において、技術上で顕在化している問題については特に指摘がなかった。

<英国を含む欧州>

2000 年代初頭より、ESFRI (European Strategy Forum on Research Infrastructures) に代表される戦略構想によって、欧州圏内における生物・医学・医療情報の共有と利活用を目指した情報通信基盤整備の重要性が指摘され、実現に向けた動きもみられている。しかしながら、具体的な推進体制、技術開発課題の展開について、資金面を含めて不備が見られる。特に、European Life Sciences Infrastructure for Biological Information (ELIXIR) における FP7 の資金援助期間 (Preparatory Phase、2007-2010 年予定) 終了後の資金繰りの目途が立っていないという報告からは、ESFRI における構想発表段階と現在では、欧州経済状況の変化により、加盟国の科学技術政策への優先意識に変化がみられていることが推察される。

<日本>

ライフサイエンス分野における大量データ時代に対応する施策として第 3 期科学技術基本計画からいくつかのプロジェクトが実行されており、主要プロジェクトである「統合データベースプロジェクト」ならびに「バイオインフォマティクス研究の推進」が、平成 23 年度より統合化されることが決まっており、現在その準備段階にある。その他、公共的なデータ提供ソースとして、バイオリソース事業が生物遺伝資源 (バイオリソース) の提供、基礎データの蓄積を含む基盤技術の開発を進めている。一方、医療分野に特化した大型プロジェクトとして、バイオバンクジャパン計画や医薬品等の安全対策の向上を高めるためのデータベース構築が検討、進捗しているが、診療情報に関する国の施策としては、内閣官房の高度情報通信ネットワーク社会推進戦略本部 (IT 戦略本部) でレセプト等の医療情報データベースについて言及されているに留まっている。個別の取り組みとしては、現在、国立情報学研究所 NAREGI (National Research Grid Initiative) や、理研サイネス (Scientists' Networking System) による「医学サイネス」(基礎研究成果データベースと臨床データベースを連結させる仕組みの構築事業) への試みが挙げられる。

これらの調査結果から、以下の 5 点を考慮して THIB の推進戦略の具体化を推し進めるべきとの知見を得た。

- 国家レベルでの医療情報と基礎生物学研究の成果情報を結びつけた大規模かつオープンなデータベースの構築と活用に関しては、世界的に重要性が認識され、インフラ構築が期待されているものの、基盤技術整備が追い付いていないのが共通の現状である。したがって、我が国において、THIB の開発を推進することで、世界をリードするものとなりうる。
- 特に、情報基盤整備に有用な要素技術の研究開発を積極的に行い、各国の、あるいは多国間の情報通信基盤整備の需要に合わせて国際展開することで、当該分野への貢献が期待できる。
- 上記国際展開に関しては、THIB を情報通信外交戦略の一環と位置付け、適切な需要をもつ国ならびに機関に関する事前の市場調査を綿密に行い、売り込んでいくことが重要である。
- すでに他の先進国でも取り組まれているように、医療サービスの安全性、有効性、

効率性の向上に向けて、診療内容やその成果をデータベースとして蓄積し、医薬品の開発や医療提供の改善に向けた情報の利活用を推進する。

- THIB の有効性を高めるには、疫学等の医療情報の蓄積が生命科学分野の基礎研究に対する新たな問題提起、成果の還元につながる仕組みについてより一層の工夫が必要である。

本ベンチマーク調査の成果は、戦略プログラム『生命・医学・医療・健康をつなぐ情報を循環させる技術と基盤の構築と活用 ～トランスレーショナル・ヘルスインフォマティクス・ベースの展開～』において提唱する研究開発戦略の具体化に役立てられる。

Executive Summary

It has been discussed that we should perform “Health Research” which enables us to make use of the fruits of latest life science research and to return it to the society as a more advanced treatment method, medicine or medical instrument. To this end, the Health Research Promotion Council was established in Japan under the control of the Cabinet Office in 2008. In the “Health Research Promotion Strategy” proposed by the Council, it is recommended that we establish “the circulation of knowledge” loop in the medical science and clinical services by clarifying the mechanism of life phenomenon and diseases, converting the clarified mechanism into diagnostic and treatment methods, identifying new research subjects through assessing the effectiveness of the methods, and carrying out research on the identified subjects.

Based on these trends, JST-CRDS provided Strategic Program entitled “Development of Translational Health Informatics Base” , in which the importance of information technology infrastructure is pointed out for the establishment of knowledge circulation structure across clinics and research laboratories. Furthermore, JST-CRDS executed international benchmarking survey to add more powerful and concrete promotion framework to the Strategic Plan by comparing the provided research and development strategy with those have been partly worked in the United States and European countries. In the benchmark survey, we analyzed following points; 1) current states of collection, storage, and extraction of large scale data in both life- and health-science fields, 2) incentive strategy to motivate the data providers, and 3) technical and systematic issues to be solved for better use of clinical informatics in basic and medical sciences. We also carried out the fact-finding mission with eminent researchers who have world-wide networks in the related fields to investigate clinical informatics in the United States and integrative research project across clinical- and bio-informatics and its technological infrastructure in European Union countries (Germany and United Kingdom). As a result, following facts were found and confirmed.

<United States>

Health Informatics industry is growing in the U.S. Insurance companies and IT companies are key players in this industry, and they provide government agencies, health care providers, research institutes, and employers with large-scale health care data. Because health care cost turns to be a burden for both private and public sector, there is strong incentive to utilize such data for improving the efficiency and effectiveness of care shared across the health care field. Regarding privacy issues around health care data, not only privacy protection but meaningful data utilization is aimed in regulations. There is no problem pointed in our discussion related to

technical aspect of the database development.

<European Union>

Importance of pan-European information technology infrastructure has been pointed out by ESFRI (European Strategy Forum on Research Infrastructures), and several related projects have already started. However, systematic flaw including financial cooperation was recognized as the bottle neck to progress the framework plan and technical development. The observation suggested that the recent deterioration in economic conditions caused the various reaction toward the priority of science and technology policy across the affiliated countries.

<Japan>

Since the execution of Third-Basic Program for Science and Technology, several projects have been worked for the large-scale data management in life science field. FY 2010 is the preparatory periods for the integration of two major projects, “Basic Program for Science and Technology (funded by Ministry of Education, Culture, Sports, Science and Technology, MEXT)” and “Institute for Bioinformatics Research and Development (BIRD) (funded by JST)” , which is expected in FY2011. There are other individual projects related to public data source, large-scale clinical informatics, and government agency research program, it is still on the process to big achievement.

Taken together with whole results, JST-CRDS would provide following strategies for the implementation of the “Development of Translational Health Informatics Base”.

- Even there would be strong consensus and expectation for the establishment and usage of nation-wide large scale and open access database linking clinical and bioinformatics, buildup of fundamental technologies is not sufficient in common across the World. Therefore, the JST-CRDS’ s strategic program of THIB is highly enough motivated and prompt investment would bring national interests to Japan.
- It is highly expected for international contribution by Japanese technology that the global promotion of element technology development based on the current demand of information technology infrastructure development in each country or across countries.
- For the above global promotion, one should rank the THIB as a part of global strategy in information technology and execute careful feasibility survey to search the market in target countries and institutions.
- In order to improve the validity of THIB, more ingenious attempt would be needed to establish the framework that clinical informatics including epidemiology data become more applicable for basic researchers with novel scientific questions.

The result of the current benchmarking survey would be used for externalization of the research and development strategic plan provided by “Development of Translational Health Informatics Base”.

目 次

エグゼクティブサマリー Executive Summary

1. 背景と目的	1
2. 調査方法	3
3. 調査のまとめ	5
4. 調査メンバーの総合所見	
4.1. 豊田 哲郎（理化学研究所 生命情報基盤研究部門）	7
4.2. 堀川 一樹（北海道大学 ニコンイメージングセンター）	9
4.3. 興梠 貴英（東京大学 健康医科学創造講座）	11
＜コラム＞ 大規模データベースを用いた医療サービスの検証例	14

<Appendix>

A1. インタビュー用資料	19
A2. 世界のバイオ／ヘルスインフォマティクススペースに関する比較調査結果	
A2.1. 要素技術研究の国際動向（学術研究）	23
A2.2. 要素技術開発の国際動向（推進政策、研究資金）	25
A3. サイトリポート	
A3.1. 米国における関係機関担当者へのインタビュー	
Agency for Healthcare Research and Quality	31
トムソン・ロイター社	34
ユナイテッド・ヘルス・グループ	38
スタンフォード大学	40
A3.2. 欧州における関係機関担当者へのインタビュー	
EMBL Heidelberg	41
CASIMIR	51
EBI／ELIXIR	54
A4. 国内の状況	
A4.1. 国内の関連施策	61
A4.2. インタビュー対象者から挙げられた国内の主要研究者	62

1. 背景と目的

優れたライフサイエンスの成果を活用し、新しい治療法や医薬品・医療機器等として社会に還元していくための「健康研究（Health Research）」（橋渡し研究・臨床研究）の必要性が謳われている。我が国においては、2008年に設立された健康研究推進会議（内閣府）が「健康研究推進戦略」を発表し、基礎研究により生命現象と疾病のメカニズムを解明し、それを診断や治療法に転換して医療を実践し、その医療効果を評価して新たな課題を設定し、それを再び基礎研究につなげていくという、医学・医療における知の循環の確立を目指すべきであるという問題意識が示されている。

上記を踏まえ、独立行政法人科学技術振興機構 研究開発戦略センター（JST-CRDS）では、戦略プログラム『生命・医学・医療・健康をつなぐ情報を循環させる技術と基盤の構築と活用 ～トランスレーショナル・ヘルスインフォマティクス・ベースの展開～』を通して、医学・医療における「知の循環」の確立に必要な情報を循環させる技術と基盤（ここでは THIB：トランスレーショナル・ヘルスインフォマティクス・ベースと名づけた）の構築と活用を提案した。THIB を構築することで、ライフサイエンス研究の成果と個々人の医療情報・診療データの相互利用が可能となり、基礎医学の進展や医療の高度化（個別化医療の実現など）に加えて、疫学研究の進展、医療プロセスの高品質化、さらには経済面からの医療の効率化といった効果が期待できる。

本戦略プログラム策定においては、有識者インタビューならびに科学技術未来戦略ワークショップ『トランスレーショナル・インフォマティクス・ベースの展開』（2010年1月18日開催）を通じた、当該分野の専門家による研究開発と推進体制に関する意見、要望を参考とした。具体的には、ライフサイエンスの最先端の知と医療を結びつけるための情報基盤、さらには臨床データを有効に利用するための情報基盤、診療データ、健康診断データ、疫学調査データなどの活用も視野にいたれた情報基盤を総合的に THIB として構築する。THIB の構築・活用に当たっては、インフォマティクスに関連する先進的研究への投資だけでなく、標準化活動や共用ツールなどのインフラ構築部分への投資と恒久的運用が必要である。また、THIB 構築の具体的推進方法としては、当初から統一的大規模システム開発を目指すのではなく、拠点を決めスモールプロジェクトでスタートさせ、インフラ構築を進めつつコミュニティを拡げていくこととし、既存システムの連携（フェデレーション）に重点を置いたシステム構築を目指す。

今回、上記提案内容に類似した取り組みを先んじて行っている欧米の組織と比較し、より強力な研究開発推進体制の具体化戦略の提案を行なうために、国際ベンチマーク調査を実施した。本国際ベンチマークでは、①ライフサイエンス、ヘルスサイエンス双方における大容量データの収集、保管、抽出技術の現状、②データ収集における提供者側へのインセンティブの付与戦略、③医療情報を基礎研究あるいは臨床現場に活用するための技術的制度的課題の克服、について、関係スタッフによる調査分析を行ったほか、当該分野の研究に詳しく、且つ海外の研究者コミュニティとの幅広いネットワークを持つ専門家に国際ベンチマーキング参加を依頼し、特に、米国における医療情報の管理、活用形態、ならび

にドイツと英国におけるとバイオインフォマティクスとヘルスサイエンスの融合研究プロジェクトと情報インフラ整備の現地調査を通じて我が国との比較を行った。

(参考)

- 戦略プログラム『生命・医学・医療・健康をつなぐ情報を循環させる技術と基盤の構築と活用 ～トランスレーショナル・ヘルスインフォマティクス・ベースの展開～』
CRDS-FY2009-SP-08
- 科学技術未来戦略ワークショップ報告書 『トランスレーショナル・インフォマティクス・ベースの展開～生命・医学・医療・健康をつなぐ情報を循環させる技術と基盤の構築と活用 ～』 CRDS-FY2009-WR-08

2. 調査方法

本調査は、JST-CRDS のメタインフォマティクスチームが、興梠貴英氏（東京大学 健康医科学創造講座）、豊田哲郎氏（理化学研究所 生命情報基盤研究部門）、堀川一樹氏（北海道大学 ニコンイメージングセンター）らの協力を得て企画立案し、各種バイオ/ヘルスインフォマティクス基盤に関連する多角的なデータに基づく国際比較解析と現地調査（米国、ドイツ、ならびに英国の研究者、バイオ/ヘルスインフォマティクス基盤担当者、科学政策担当者へのインタビュー）から構成されている。国際比較解析は世界分布、標準化技術開発プログラム、発表論文データ、テクニカルレポート等のデータ分析に基づいて実施した。また、米国および英国における現地調査については、下記の要領で実施した。

訪問先の選定： 本調査においては、米国では大規模医療情報の収集と分析、それに基づくサービスの具体例、それらを支える基盤技術開発と基礎研究者への展開戦略に注目し、欧州では、基礎研究において得られた大規模データの管理に有用な基盤技術・管理体制・人材育成と、医療分野への展開戦略に注目して、訪問期間を選定した。特に、大容量のデータを扱う機関における情報インフラの運営体制が充実しているか、そして、データの利用者側、提供者側、双方の立場からの意見が組織運営やデータの利活用に活かされているか、という観点を重視した。

米国の訪問先として、大規模医療情報を集計、解析して、医療サービスの評価や改善、再構築などを担当する行政機関である米国医療の質研究機構（Agency for Healthcare Research and Quality, AHRQ）、同機構と共同で情報を収集し、かつ患者向けの医療サービスとして活用している保険会社の United Health（ユナイテッド・ヘルス）社およびグループ企業、さらに大規模医療情報データベースを管理、運用する Thomson Reuters（トムソン・ロイター）社、医学研究と医療情報の橋渡しを実際に行なっている Stanford 大学を選定した。欧州においては、生物学分野における大容量データの収集・分析による実験成果を数多く挙げている European Molecular Biology Laboratory (EMBL) 本部（ドイツ・ハイデルベルグ）と、FP7 において欧州圏全体のライフサイエンス分野のインフォマティクス基盤の連携強化を目指しているプロジェクト European Life sciences Infrastructure for Biological Information (ELIXIR) の本部である、EMBL の下部組織 European Bioinformatics Institute（英国・ヒンクストン）、ならびにマウスに特化したバイオインフォマティクス研究者の国際連携組織である Coordination and Sustainability of International Mouse Informatics Resources (CASIMIR) を選定した。

インタビュー実施要領： JST-CRDS の設立経緯とミッション、本調査の目的の説明資料とともに個々の訪問先に応じた質問票を作成し、事前に送付した（送付資料については付録参照）。訪問当日は質問票に沿ってインタビューを実施した。インタビュー記録は調査者全員で内容確認を行なった。以下に、米国、ドイツ、英国におけるベンチマーク調査の実施要領を記す。

<米国ベンチマーク調査>

Agency for Healthcare Research and Quality

調査日：平成 22 年 3 月 23 日

調査者：山本 雄士（CRDS 臨床医学ユニット フェロー）、興梠 貴英（東京大学 健康医科学創造講座）

トムソン・ロイター社

調査日：平成 22 年 3 月 24 日

調査者：山本 雄士（CRDS 臨床医学ユニット フェロー）、興梠 貴英（東京大学 健康医科学創造講座）

ユナイテッド・ヘルス・グループ

調査日：平成 22 年 3 月 25 日

調査者：山本 雄士（CRDS 臨床医学ユニット フェロー）、興梠 貴英（東京大学 健康医科学創造講座）

スタンフォード大学

調査日：平成 22 年 3 月 26 日

調査者：山本 雄士（CRDS 臨床医学ユニット フェロー）、興梠 貴英（東京大学 健康医科学創造講座）

<ドイツベンチマーク調査>

European Molecular Biology Laboratory (EMBL) Heidelberg

調査日：平成 22 年 3 月 29 日 – 平成 22 年 3 月 30 日

調査者：福士 珠美（CRDS ライフサイエンスユニット フェロー）、豊田 哲郎（理化学研究所 生命情報基盤研究部門）、堀川 一樹（北海道大学 ニコンイメージングセンター）

<英国ベンチマーク調査>

Coordination and Sustainability of International Mouse Informatics Resources (CASIMIR)

調査日：平成 22 年 3 月 31 日

調査者：福士 珠美（CRDS ライフサイエンスユニット フェロー）、豊田 哲郎（理化学研究所 生命情報基盤研究部門）、堀川 一樹（北海道大学 ニコンイメージングセンター）

European Bioinformatics Institute (EBI) / European Life sciences Infrastructure for Biological Information (ELIXIR)

調査日：平成 22 年 3 月 31 日

調査者：福士 珠美（CRDS ライフサイエンスユニット フェロー）、豊田 哲郎（理化学研究所 生命情報基盤研究部門）、堀川 一樹（北海道大学 ニコンイメージングセンター）

3. 調査のまとめ

今回、バイオインフォマティクス、ヘルスインフォマティクスの技術基盤と研究開発動向の国際比較調査を実施した結果、医療情報と基礎生物学研究の成果情報が総合的に結びついて大規模（国、州などのレベル）かつオープンに展開されている事例はないことが明らかになった。ただし、国によっては医療情報の社会還元スキーム（医療機関、製薬企業、一般市民への情報公開など）や研究機関内での小規模でクローズドな情報循環の例が見られ、国際連携による情報基盤整備の重要性も認識はされていることが確認できた。その他、米国、英国を含む欧州、日本において、それぞれ以下のような動向が明らかになった。

<米国>

米国では、大規模な医療情報データベースを基盤とした産業が発達している。こうした産業の主要なプレイヤーとしては保険会社、IT企業の双方があり、これらが行政や医療機関、さらには研究機関や国民向けに情報を基盤としたサービスを提供している。わが国とは異なり、米国の医療制度では民間企業が医療保険財政にも大きく関わっているため、大規模医療情報を用いてより効率的、効果的な医療サービスを実現する、というニーズが産および官で共有されている印象が強い。あわせて、医療情報の扱いについても、個人情報の保護の観点とその有意義な利活用の観点の両面から規制が整備されている。また、これら情報基盤の構築において、技術上顕在化している問題については特に指摘がなかった。

<英国を含む欧州>

2000年代初頭より、ESFRI (European Strategy Forum on Research Infrastructures) に代表される戦略構想によって、欧州圏内における生物・医学・医療情報の共有と利活用を目指した情報通信基盤整備の重要性が指摘され、実現に向けた動きもみられている。しかしながら、具体的な推進体制、技術開発課題の展開について、資金面を含めて不備が見られる。特に、European Life sciences Infrastructure for Biological Information (ELIXIR) における FP7 の資金援助期間 (Preparatory Phase、2007-2010 年予定) 終了後の資金繰りの目途が立っていないという報告からは、ESFRI における構想発表段階と現在では、欧州経済状況の変化により、加盟国の科学技術政策への優先意識に変化がみられていることが推察される。

<日本>

ライフサイエンス分野における大量データ時代に対応する施策として第3期科学技術基本計画からいくつかのプロジェクトが実行されており、主要プロジェクトである「統合データベースプロジェクト」ならびに「バイオインフォマティクス研究の推進」が、平成23年度より統合化されることが決まっており、現在その準備段階にある。その他、公共的なデータ提供ソースとして、バイオリソース事業が生物遺伝資源（バイオリソース）の提供、基礎データの蓄積を含む基盤技術の開発を進めている。一方、医療分野に特化した大型プロジェクトとして、バイオバンクジャパン計画や医薬品等の安全対策の向上を高めるためのデータベース構築が検討、進捗しているが、診療情報に関する国の施策としては、

内閣官房の高度情報通信ネットワーク社会推進戦略本部（IT 戦略本部）でレセプト等の医療情報データベースについて言及されているに留まっている。個別の取り組みとしては、現在、国立情報学研究所 NAREGI（National Research Grid Initiative）や、理研サイネス（Scientists' Networking System）による「医学サイネス」（基礎研究成果データベースと臨床データベースを連結させる仕組みの構築事業）への試みが挙げられる。

これらの調査結果、そして我が国の関連動向を踏まえ、以下の 5 点を考慮して THIB の推進戦略の具体化を推し進めるべきとの知見を得た。

- 国家レベルでの医療情報と基礎生物科学研究の成果情報を結びつけた大規模かつオープンなデータベースの構築と活用に関しては、世界的に重要性が認識され、インフラ構築が期待されているものの、基盤技術整備が追いついていないのが共通の現状である。したがって、我が国における関連分野への技術開発の重要性の認識と、推進情の課題認識を解決する THIB 研究開発戦略は、国際競争に参入するに十分な高水準にあり、本研究開発への迅速な投資は国益を生むと考えられる。
- 特に、情報基盤整備に有用な要素技術の研究開発を積極的に行い、各国の、あるいは多国間の情報通信基盤整備の需要に合わせて国際展開することで、当該分野への貢献が期待できる。
- 上記国際展開に関しては、THIB を情報通信外交戦略の一環と位置付け、適切な需要をもつ国ならびに機関に関する事前の市場調査を綿密に行い、売り込んでいくことが重要である。
- すでに他の先進国でも取り組まれているように、医療サービスの安全性、有効性、効率性の向上に向けて、診療内容やその実績をデータベースとして蓄積、公開し、新たな医療技術の開発や医療提供の質の評価・改善に向けた情報の利活用を推進することが重要である。
- THIB の有効性を高めるには、疫学等の医療情報の蓄積が生命科学分野の基礎研究に対する新たな問題提起、成果の還元につながる仕組みについてより一層の工夫が必要である。

4. 調査メンバーの総合所見

4.1. 豊田 哲郎（理化学研究所 生命情報基盤研究部門）

2010年3-4月の欧州の調査ではドイツおよび英国でデータドリブンの研究活動を行う様々な研究者と意見交換を行う中で、データシェアとツールシェアに関する現状と意識を感じ取ることができたことは大変有意義であった。私自身も普段から様々な分野における世界のデータベース関係者と接しており、そこでの経験と、今回の調査での経験が結びつくことで見えてきた総合所見を、トランスレーショナルリサーチ分野発展のための提言とあわせて報告させていただく。

データシェアを介した連携では、「データ提供者 A」と「データ受領者 B」の二者だけでなく、「受益者 C」、および「資金提供者 D」を加えた少なくとも四者のインセンティブを分析の対象とする必要があることを一連の意見交換を通じて実感した。これら四者のインセンティブが明確化されていないケースでは継続的なデータシェアの関係がうまく形成できていない。

例えば、多くの研究系データベースでは、A, B, C が全て研究者コミュニティでクローズしており、その効果が国民 (D) から実感されにくい。このため、国民から見て明確な資産形成がなされるのがわかりやすい初期投資のフェーズ（リソース形成、システム開発）ではファンディングを受けやすいのに対し、その後の継続的な維持費についてはそのコストに見合った恩恵を国民が実感できないため、各国とも共通して維持費のねん出に苦労している。CASIMIR や ELIXIR はこの維持費を獲得するための活動であるが、資金提供者 (D) を説得するための苦労がうかがえる。日本の統合データベース事業でも同様の苦労がみられる。以上を踏まえると、受益者 (C) の中に国民を加えていく努力が重要であることがわかる。

トランスレーショナルリサーチでは、医療現場と基礎研究との間で知の循環を確立する必要がある、同一人物が医療と基礎研究の両方に携わるケースを除外すれば、医療現場の人物 (A) と基礎研究現場の人物 (B) との間を繋ぐデータシェアの関係を確立し、有意義な連携を持続させることが課題となる。上記の考察からトランスレーショナルリサーチの受益者 C として、研究者コミュニティだけではなく国民も加えていく必要があるが、幸いなことに、医学・医療分野では“患者”が受益者 C となりえるので、インセンティブデザインは他の研究系データベースよりも容易であると考えられる。

一般的に患者は、より高い医療満足度を得るために、ネットを使って情報を取得しようとする傾向がある。例えば米国では PatientsLikeMe.com というサイトがあり、一般の人が自分の健康状態をネットで積極的に公開し、自分と似た症状の患者を見つけ出して情報を交換し合っている。日本でもネットを使って患者同士の情報交換を始めている（例えば、<http://www.tobyo.jp/>）。インターネットが登場する前は、医師やマスコミから提供される医療情報を国民は受け取るだけの一方向型の情報発信であったが、インターネット登場後は、あらゆる分野で国民のだれでも自由に情報交換ができるようになり、その情報革命が医療分野にも波及しつつある。

そこで、我が国における理想的なトランスレーショナル・ヘルスインフォマティクス・ベース (THIB) では、臨床と基礎の研究者 A, B だけでなく一般国民である患者 C にとっ

でも有用なしくみでなくてはならない。患者 C へのよりよい医療サービスの提供をインセンティブとして、臨床と基礎の研究者 A,B の間の連携が促進されるためには、患者 C に対してトランスレーショナルリサーチの存在とその効果・リスクが正しく伝えられ、患者 C の希望によりその主治医 A がトランスレーショナルリサーチの専門医（基礎研究者 B）にコンタクトすることで、A と B が積極的に協力しあって、より良い医療サービスを行おうとする仕組みが必要である。このような A,B,C の連携モデルでトランスレーショナルリサーチを駆動する総合的なシステムを日本版 THIB として立ち上げることが喫緊の課題である。

その原型となるモデルとして、Primary Immunodeficiency Database in Japan (PIDJ、<http://pidj.rcai.riken.jp/>) による原発性免疫不全症分野の専門医 (B) と臨床医 (A) を繋ぐネットワークがある。これは、日本型の優れたモデルとして国際的にも評価され、財団 (D) である Jeffrey Modell Foundation (JMF) からのファンディングを受けている。現在の PIDJ モデルでは、患者 C への情報提供は疾患ごとに分かれた患者の会による電話相談などを中心に行われているが、THIB ではこのようなモデルを他の疾患分野にも拡大しつつ、患者のために疾患分野横断的に集約化してインターネットから患者に情報提供を行うことが望ましい。例えば、THIB では、患者が自らの症状や体験をキーワードにつかって容易に専門医を探し出せる仕組みがあり、そこで提供される「電子紹介状」を患者 C が自ら印刷して主治医 A に手渡すことで、これがきっかけとなり、主治医 A が THIB にアクセスして専門医 B にコンタクトできるシステムも考えられる。こうした悉皆的なしくみはクラウドを活用した大規模かつセキュアなシステムが必要であり、そのような THIB の開発によって国内およびアジアのハブとしての役割で日本が優位に国際貢献できるチャンスがある。そのような THIB では、日本語と英語の両言語をサポートしつつ、日本の国民性や法律に矛盾しないものが必要であり、米国型の電子カルテの存在を前提としたシステムではなく、電子カルテのない小規模の診療所の医師 (A)、基礎研究にも詳しい専門医 (B)、患者 (C)、ファンディングエージェンシー (D) をインターネットで繋ぐ必要がある。このようなシステムは、日本だけでなくアジア諸国へのハブ機能を提供することに役立つであろう。

多分野多国間で同一の THIB を共有することは「ツールシェア」であるが、ツールシェアはデータシェアよりも実現しやすく経済効果が明確なため進みやすい。例えば EMBL の細胞イメージング分野では、測定機器などのツールシェアは進んでいた。EMBL ではデータシェアはそれほど進んでいなかったが、ツールシェアでは計測器以外に、データストレージや計算機のシェアも行っていた。すなわち、個々人のデータがセキュアに保護され、他者が勝手にアクセスできない範囲で、それ以外のツールを他からのサービスとして依存し、シェアする傾向にある。このため、クラウド化によりデータシステムを他に依存しつつも、自分の研究データに十分なセキュリティをかけることができれば、クラウドシステムをツールとして共有利用するインセンティブは働くと思われる。

以上の要件を満たす要素技術の開発とフィージビリティテストを早期に実施することで、日本発の連携モデルと情報インフラを立ち上げてこの分野の発展に貢献できるものと期待できる。

4.2. 堀川 一樹（北海道大学 ニコンイメージングセンター）

JST-CRDS「トランスレーショナル・ヘルスインフォマティクス・ベース (THIB)」国際ベンチマーク調査のヨーロッパ動向調査に関する所見を述べる。訪問先である EMBL、CASIMIR、EBI-ELIXIR そして EBI Data Center のいずれにおいてもトランスレーショナルリサーチは最重要視されている戦略目標であり、我々と問題意識を共有していた。以下、実際のインタビューにより得られた、ヨーロッパにおけるトランスレーショナルリサーチの現状と問題点を、特にインフラ整備、人材育成そして画像情報への対応に整理して要約する。

1. インフラ整備

EBI Data Center と隣接する Sanger Institute では膨大な塩基配列情報を保管・処理できるデータセンターが稼働しているのは周知の事実である。現在の処理能力もさることながら、今後指数関数的に増大すると予測される大量のデータに対応できるだけのシステムの拡張性が保証されていたことは注目に値する。プロセッサファームと呼ばれる数千の CPU とテラバイトハードディスクの集合体は、その汎用性、保守の容易さ、費用対効果などあらゆる観点から合理性に富んでおり、国内においてもスパコン等と併用されるべき強力なインフラであろう。

2. 人材育成

EBI-ELIXIR は次世代バイオインフォマティクスの情報基盤拠点である。2011 年からの本格運用を前に、2007-2010 年の準備期間中に十分な運営費が投入され、専任のバイオインフォマティシャンならびにアドミニストレーターが動員されている。持続的なインフラ整備と人材育成を可能にするこうしたやり方は、長期プロジェクトを押し進める際に非常に強力な方法であることを実感した。これは国内のトランスレーショナルリサーチが直面する状況と大きく異なっており、今後の国際連携を優位に進めるためにも参考にすべき方法の一つである。

3. 画像データへの対応

CASIMIR は、疾患モデルマウス・フェノームデータベースのヨーロッパ拠点で、既にアメリカ、カナダそして日本を含めた国際連携が図られている。変異マウスの表現型解析におけるオントロロジーの整備や解析プロトコルの標準化などの技術要素は一定の成功を収めており、ヒトの臨床データを扱う THIB の構築にあたっても学ぶべき点は多いと感じた。ただし、現時点では画像情報へは対応していないこと、またデータベースの拡張性について技術的なハードルが存在するなど解決されるべき課題も散見された。THIB を押し進める上では、CT や MRI といったヒト疾患画像のデータベース化も必要である。またデータベースの拡張が可能なセマンティック web 化など IT 技術が果たしうる役割は大きく、こうした技術要素の開発について日本が先導性を発揮する余地が十分にあると感じた。

EMBL では基礎ライフサイエンス分野において大容量画像データ時代の到来を確信させる研究を目の当たりとすることになった。特に、E. Stelzer らによる DSLM(digital scan light sheet microscopy) の開発と改良、ならびに R. Pepperkok と J. Ellenberg グループによる画像取得の自動化と RNAi による機能アッセイを組み合わせたハイスループットスクリーニングシステムは注目に値する。既に取得されるデータサイズは 3-40Tb/experiment に達しているが、DSLM については現在の 4 次元データから（細胞の位置と時間変化）、最大 10 次元化を目指した改良と製品化が進められており、データサイズは今後飛躍的に増えることは間違いない。すでにデータ処理が律速となっている状況で、今後確実に起こるデータ量爆発を見据えた革新的技術の開発が求められる。

なお、これらの膨大な情報の一部は、既にデータベースとして公開されている。ここではイメージライブラリーとゲノム情報がリンクされる等、有効活用に向けた基盤整備が急速に進行している。ヘルスインフォマティクスにおいても、CT や MRI による臨床画像データを発見的利用につなげるためにゲノム情報等とリンクさせた統合的なデータベースの構築が重要な課題となるので、EMBL での先行例に学ぶ点が多い。

まとめ

今回のインタビューでは、主にライフサイエンス分野でのトランスレーショナルリサーチについての戦略と得られつつある成果を目にすることができた。一方で、臨床データを対象にしたヘルスサイエンス分野におけるトランスレーショナルリサーチはまだまだ発展途上の段階にあり、日本が将来の国際連携をリードすることは可能であると感じた。既にナノテクやバイオプローブ開発などに支えられた光イメージング技術について日本は世界をリードする立場にある。また、大量データのハンドリングを可能にする高速通信網やスパコンなど、社会インフラが整備されつつあることも大きなアドバンテージである。これらの突出した技術的要素を「ガラパゴス化」させることなく有効活用するためにも、技術要素の有機的なネットワーク化を、その利用方法も含めた「パッケージ」として国内外にいち早く示す必要がある。このためにも、様々な研究分野の知を循環的に相互作用させるかけはしとして、境界領域で活躍できる人材を長期的な視点で育成・活用する仕組みを整備することが重要であろう。

4.3. 興梠 貴英（東京大学 健康医科学創造講座）

米国大規模医療情報データベースの現状について

米国においては、公的・私的な大規模医療情報データベースが複数存在している。公的なものとしては米国厚生省 (HHS) 下の医療研究・品質局 (AHRQ) のサポートのもと、連邦、州、医療業界が協力して構築している公的医療データベース群 (H-CUP) がある。これらは年々データベース規模を拡大してきており、全米 1000 以上の病院から得られた入院データベースである NIS の最新版には 40 の州から年間約 800 万件の退院時データが蓄積されている。

収集しているデータはデータベースによって異なるが、例えば NIS であれば主要疾患名、二番目に重要な疾患名、主要手技名、二番目に重要な手技名、入院時、退院時の状態、患者背景（性別、年齢、人種、地区別平均年収等）、保険の種類、支払総額、在院日数、病院の種類（所有形態、規模、教育病院かどうか等）を含む 100 以上の項目が設定されている。これらのデータは医療サービスの利用・コスト、治療法のばらつき、医療費増大、病院の財政難、治療の有効性、医療の質、医療政策変更の影響、医療へのアクセス、医療技術の伝播、特定の人口集団による医療サービス利用の分析に用いられており、申請を行って認められれば H-CUP からデータを購入することが可能である。

私企業による大規模データベースの例として調査の中でトムソン・ロイター社、Ingenix 社の事例を調査してきた。トムソン・ロイター社ではそれまで蓄積してきた医療情報データベースを MarketScan として外部に提供しており、研究目的で提供している MarketScan for research には延べ 1.13 億人のデータ、うち入院データ 4600 万件ときわめて規模が大きく、最近 5 年間で MarketScan for research を用いた研究成果は査読付き雑誌に 200 編以上発表されており、様々な分野における医療評価のための強力なツールとなっている。

米国最大級の保険者ユナイテッド・ヘルスケア社と同じユナイテッド・ヘルス・グループ傘下の Ingenix 社は償還請求データ、処方箋データを中心とした大規模データベースを構築している。これまでに延べ 7000 万人の医療データ、10 年間連続して追跡できる人数だけでも 1500 万人分とこちらも非常に大きなデータベースであり、そのデータを元にヘルスケア内の各ステークホルダーに対してコンサルティングビジネスを展開している。

こうしたデータベースを構築・分析することによって、医療制度、医療費、医療の質等のマクロ的な観点からのみならず、各病院における医療コスト削減可能箇所の提示、個々の患者に対して医療サービス間の連携のギャップをどこにあるのかを指摘できるようになる等幅広い分野にわたって改善を提示できるようになっている、ということが分かった。

大規模医療情報データベースの必要性について

我が国では財政が逼迫している中でも今後ますます人口の高齢化は進んでいく。このときに重要となってくるのはどこに重点的にリソース配分すべきなのかという判断と、配分されたリソースが有効に使われているのかということの緻密な検証である。欧米の研究でも示されているように、大規模データベースはそのためにはきわめて有効であり、今後日本においてもきちんとしたデータベースを構築する必要がある。

我が国において大規模医療情報データベースを構築する上での現状と課題

我が国でも 2011 年にレセプトのオンライン請求が原則義務化され、請求データを基礎とした大規模データベースを構築しうる環境が整う予定である。しかし、現在のレセプトには例えば主病名、併存病名が明示されないために、欧米における報告と同等の質で疾患との関連を分析することが困難である、という問題がある。つまり、オンラインレセプトデータを中核としたデータベースを構築するのであれば、一次データの質を高くすることが必須であり、関連する臨床情報（例えば病名）をより正確に登録する仕組みを導入する必要がある。

急性期病院で最近採用が増えている DPC では「医療資源を最も投入した傷病名」、「入院の契機となった傷病名」等重みづけをした病名が登録されるために、疾患毎のリソース消費分析などを行うことができる。データベースの有用性を高める一つの方法としては外来通院治療や出来高払いの場合でも DPC に準じた形で情報提出を義務づける、というものが考えられる。

逆に欧米のデータベースにおいても現時点では詳細な臨床情報（血液検査値、各種生理検査結果、画像検査結果）までは蓄積されていない。もし、今後日本において大規模データベースを構築する際に、十分な臨床情報も同時に蓄積していくことができれば、世界に先駆けて質が高く、より深い解析が可能なデータベースを構築することも十分可能である、と考えられる。

また調査の結果、米国において大規模医療情報データベースが構築可能な理由として法的整備（HIPAA（医療保険の相互運用性と説明責任に関する法律）の施行）、保険加入時の契約書に匿名データの利用に関する項目が含まれている、といったものがあると推測され、完全匿名化したとしても個人情報保護との関連でこうした大規模データベースにデータを登録しにくい我が国において参考にするべきものと考えられた。

我が国への提言

日本における現状およびこれまでの議論をまとめると、

- ・大規模医療情報データベースは医療の質・安全性評価・リソース評価にきわめて有効である
- ・日本においてはいかなる形でこうした大規模データベースが存在しておらず、今後高齢化がますます進み、医療資源の有効活用を行うことが必須となってくる我が国においては構築していくことが重要である
- ・データベースの質確保のためには一次データの質が重要であり、従来のレセプトデータのみでは十分な質が確保できず、従って保険請求における制度改革も含めた変革が大規模データベース構築の際には必要となる可能性が高い
- ・データは蓄積するだけでは意味がなく、解析を行って初めて意味を持つので、解析を行う視点でのデータ構築を行うことが重要である。また米国 H-CUP のようなデータベース維持・管理を行うセンターも必要となってくる
- ・現時点では米国の大規模医療情報データベースにおいても血液検査結果などの臨床情報はまだほとんどデータベースに含まれていない。今後日本で構築する際に、より臨床情報に踏み込んだデータも同時に蓄積する仕組みにすれば世界で最高水準のデータベースを構築可能である

従って、今後大規模医療情報データベースの構築に当たっては下記を提言したい

- ・保険請求に伴って医療機関が支払機関に提出する患者情報を医学・医療的により正確にするような制度改革を行う。またデータ項目数も十分に増やし、後にリソース分析を行うに足るだけの項目群とすることも必要である。
- ・大規模医療情報データベースセンターの設立を行う。ここに完全匿名化したデータを集約するとともに、広く解析目的で用いることができるような仕組みを作る。米国 H-CUP がモデルになると考えられる。
- ・こうした医療情報の取り扱いが大規模医療情報データベース構築の上で障害とならないよう、運用・管理のルール作りをきちんと行う。

コラム 大規模データベースを用いた医療サービスの検証例（興梠先生資料より）

大規模データベースを用いた具体的な研究として、例えばガイドラインに沿った医療が実際に行われているか、病院規模とアウトカムの関係はどうか、といったことが検証されている。事例を以下に示す。2010年5月号のThe American Journal of Medicineにおいて「Are Atrial Fibrillation Patients Receiving Warfarin in Accordance with Stroke Risk? (心房細動患者は脳卒中リスクに合わせたワーファリン投与をされているか?)」という題名の論文が発表された。これは、トムソン・ロイター社のMarketScanデータベースを用いて行われた研究で、実地医療がガイドライン通りに行われているのか、ということを検証している。

心房細動は不整脈の一種で、年齢とともに罹患率が上昇し、欧米では60歳で約2%、80歳で約10%に認められる比較的良好に見られるものである。心房細動そのものは多くの場合心臓の機能に大きな影響をもたらさず、症状がない限りは昔は放置されていることが多かった。しかし、これが原因で心臓の中に血栓が生じて脳梗塞の重要な原因の一つとなることが分かり、近年の各国の診療ガイドラインではそうした血栓ができないような予防治療（抗凝固療法、多くの場合ワーファリンの継続的服用）を行うことが推奨されている。ただし、全ての心房細動患者が等しく血栓を生じ、脳卒中を起こすリスクをかかえているわけではなく、高血圧、糖尿病を合併しているか等によってリスクレベルが異なる。こうしたリスクレベルを評価するスコアとしてCHADS₂スコアというものが開発され、現在日本のガイドラインでは1点であれば抗凝固療法を開始してもよく、2点以上は禁忌がない限りは抗凝固療法を開始すべきである、とされている。

論文ではMarketScanデータベースから2003年1月1日～2007年9月30日の期間で一度でも心房細動の診断をされた573,519人の患者を抽出し、さらに解析対象としてふさわしい条件をかけていて以前から心房細動を有していた119,486人、新たに心房細動を発症した51,907人を同定している。これらの患者のCHADS₂スコア分布を図1に示す。

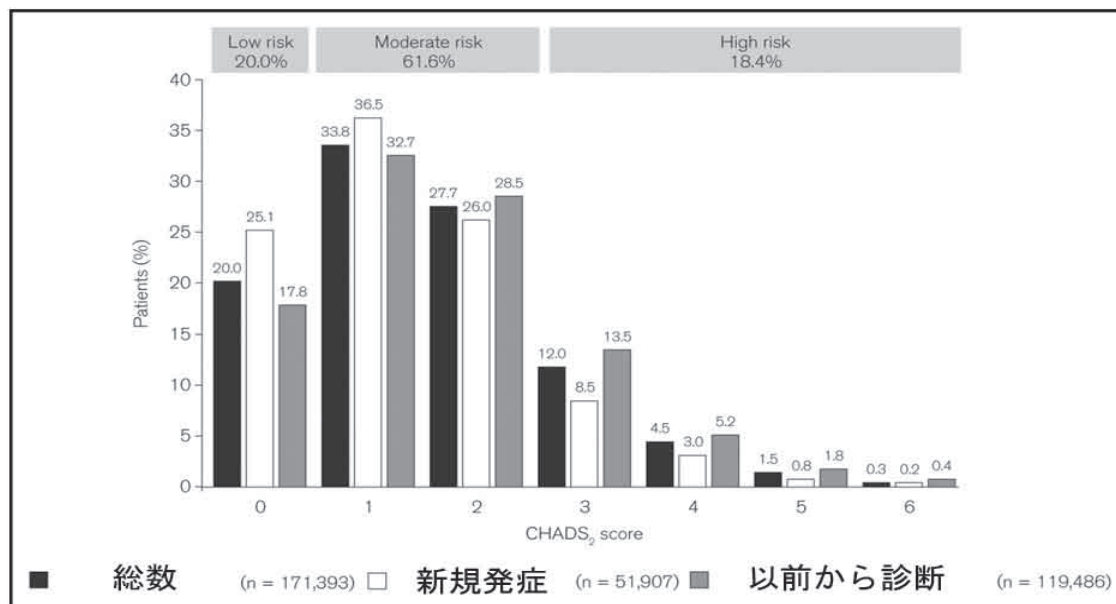


図1 CHADS₂スコアに基づいた患者数の分布

ここではスコア0点を低リスク群、1～2点を中等度リスク群、3点以上を高度リスク群に分けており、それぞれ総患者数の20%、61.6%、18.4%を占めている。

これらの患者が実際に抗凝固療法（ワーファリンの投与）を受けていた割合を計算した結果を図2に示す。

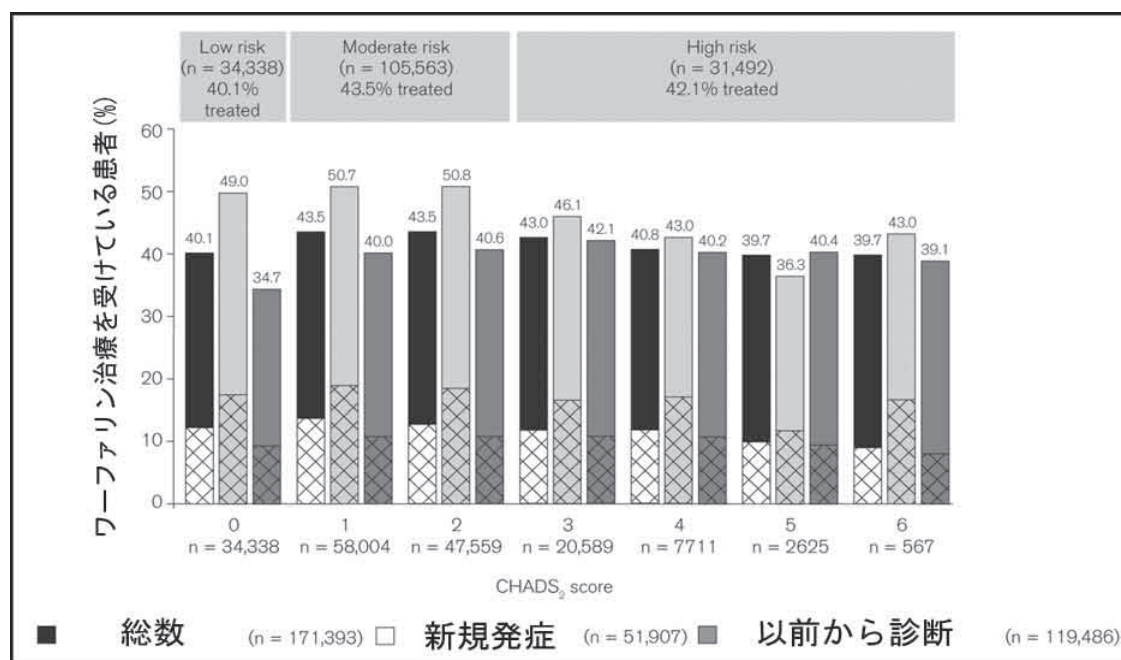


図2 診断後30日以内のワーファリン投与割合、網掛けは投与開始後180日の時点で連続投与を受けていた患者の割合

これを見ると分かるように、いずれのリスクレベルでもワーファリンを実際に投与開始されたのが半数以下でしかなく、またリスクレベル間で差がなかった。さらにその中で半年間治療が継続されたのは32.8%（総数からの割合だと13.8%）に過ぎなかった。このことから、筆者らは実地においては十分ガイドライン通りの医療が行われておらず、高リスク患者群においてはワーファリン服用を妨げる何らかの要因が存在する可能性はあるがひとまず新たなエビデンスが得られるまでは総合的にリスクを評価しつつガイドラインに従ったワーファリン使用を心がけるべきだ、と結論づけている。

この研究結果からは著者らの結論とともに、実地医療を反映した大規模データベースから得られた知見からならではの貴重な示唆が得られている。

- ・大規模臨床試験や疫学調査を通じてエビデンスが構築され、ガイドラインが制定されても実際に遵守される割合は低い、ということが分かったことはガイドラインを制定した後の実地医療評価がデータベースを解析することによって可能となっているということである。
- ・また実際にガイドラインに関する正確かつ具体的な情報が現場に浸透していないのであれば、そこをターゲットにリソースを投入することによって医療の質向上が図られるかもしれない。
- ・逆にガイドラインと実地医療との乖離が単に医師の怠慢・無知のためではなく、ガイドラインには収まりきらない現場における事情が存在するのではないか、ということも考えられる。従ってガイドラインと実地医療との乖離が起こる原因を探求することによってよりきめ細かいガイドラインを制定できるのではないか、ということも考えられ、こうした解析をフィードバックすることによって将来的に医療の質向上に寄与できる可能性が示唆される。

次に、医療機関の規模と入院後30日時点での死亡率の関係を米国 Medicare の大規模データを用いて、比較的ありふれた疾患である肺炎、心筋梗塞、心不全を対象に解析した報告を紹介する。報告は2010年3月25日号の New England Journal of Medicine に掲載されたものである。

研究は2004年1月1日から2006年12月31日までに保険として Medicare を用いた65歳以上の入院患者を対象としている。データ抽出の結果、3年間で4,128病院において734,972件の急性心筋梗塞による入院、4,679病院において1,324,287件の心不全による入院、4,673病院において

1,418,252 件の肺炎による入院を同定している。これらの入院患者の入院後 30 日時点の死亡率と病院の規模の関係を、病院の性格や患者の年齢、性別、合併症等も考慮して解析した結果、図 3 のような関係があることが分かった。いずれの疾患においてもある程度規模が大きくなってくると大きな違いがなくなり、急性心筋梗塞では年間 610 人以上を治療している病院間では差がなくなり、同様に心不全では 500 人以上、肺炎では 210 人以上では差がなくなる。しかし、規模が小さくなるに従って入院後 30 日時点での死亡率は高くなり、従来心臓カテーテル治療や心臓外科手術で報告されていた診療患者数とアウトカムの関係がこうした内科疾患においても存在することが示された。

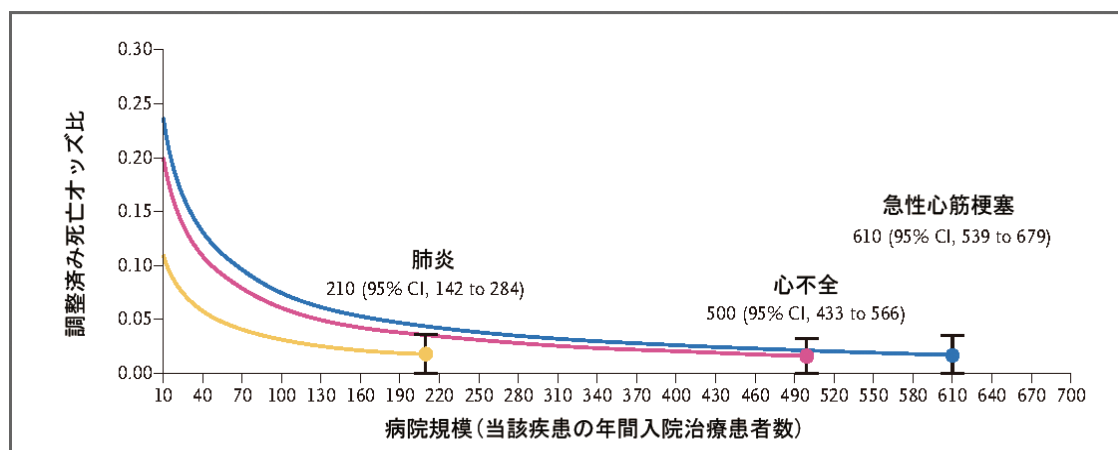


図 3 病院規模と調整済み死亡オッズ比の関係

日本においても、欧米に見習って医療機関を集約化してセンター化すべきだ、という議論が以前よりあるが、医療制度の異なる日本において実際に様々な疾患で規模とアウトカムにはどのような関連があるのか、ということについてはデータがないためにほとんど検証がされていない。

これらの報告から、大規模データベースを構築し、解析することが、医療の質の評価・向上につながる事が分かる。

<Appendix>

A1. インタビュー用資料

Center for Research and Development Strategy

International benchmark on technology levels of Translational Healthinformatics Base Research & Development

Japan Science and Technology Agency (JST)
Center for Research and Development Strategy (CRDS)

March, 2010

Copyright©2010 Center for Research and Development Strategy. All Rights Reserved.

Center for Research and Development Strategy

In this presentation we will explain...

- What is the JST and CRDS
- What is international benchmark
- Questionnaires for each researcher/institute

URLs
Japan Science and Technology Agency: <http://www.jst.go.jp/EN/index.html>
Center for Research and Development Strategy (CRDS): <http://www.jst.go.jp/EN/menu1/crds.html>

Copyright©2010 Center for Research and Development Strategy. All Rights Reserved.

Center for Research and Development Strategy

What is JST?

As an organization which plays a central role in the implementation of the Science and Technology Basic Plan, Japan Science and Technology Agency (JST) comprehensively promotes all of the process from the creation of knowledge, which leads to innovation in Japan, to the return of research results to the society/nation.

As the bases of such process, JST also provides scientific and technical information, increases the understanding of science and technology (S&T), and promotes strategic international activities.

History

Japan Science and Technology Agency (JST) was established in October 2003, succeeding the Japan Science and Technology Corporation (JST Corp.) which was established in September 1999. The JST Corp. was established by the merger of the Japan International Center of Science and Technology (JICST) and the Research Development Corporation of Japan (RDC).

http://www.jst.go.jp/EN/jstguide2007.pdf

Copyright©2010 Center for Research and Development Strategy. All Rights Reserved.

Center for Research and Development Strategy

Structure of JST

Current President
Koichi Kitazawa Ph.D.

No. of Permanent Staff: about 1,500
No. of Researchers and Others Employed by JST: about 2,500

http://www.jst.go.jp/EN/organization/index.html

Copyright©2010 Center for Research and Development Strategy. All Rights Reserved.

Center for Research and Development Strategy

Revenue and Expenditures of JST in Fiscal 2009 (Estimate, million Yen)

Revenue: 115,870 million Yen
Expenditures: 115,376 million Yen

Transition of Yearly Operating Expenditures

Amount of Initial Budget (Japan Science and Technology Agency)
Amount of Initial Budget (Japan Science and Technology Corporation)

http://www.jst.go.jp/EN/about/funds.html

Copyright©2010 Center for Research and Development Strategy. All Rights Reserved.

Center for Research and Development Strategy

International Activity of JST

Strategic International Cooperative Program

http://www.jst.go.jp/inter/english/

Based upon intergovernmental talks and other agreements on cooperation in S&T fields, MEXT designates certain countries and fields as especially important in order for Japan and its counterpart countries to promote cooperation. On such designation, JST, with the Strategic International Cooperative Program, and its counterpart country jointly make further advances in research exchanges from both countries. Specifically, JST supports joint researches and research exchanges among research teams, dispatching and hosting researchers and organizing symposia and seminars.

http://www.jst.go.jp/EN/CorporationGuide.pdf

Copyright©2010 Center for Research and Development Strategy. All Rights Reserved.

Center for Research and Development Strategy

International Activity of JST (cont'd)

Science and Technology Research Partnership for Sustainable Development <http://www.jst.go.jp/global/english/>

JST supports international joint research cooperation between Japan and developing countries for resolving global issues such as environment/energy, natural disaster prevention, and infectious diseases control. Such research cooperation is conducted in collaboration with JICA, an organization that implements ODA technical cooperation.

http://www.jst.go.jp/EN/CorporationGuide.pdf

Copyright©2003 Center for Research and Development Strategy. All Rights Reserved.

Center for Research and Development Strategy

As of September 1, 2009

Strategic International Cooperative Program

- USA, China, India, UK, Sweden, France, Germany, South Africa, South Korea, Denmark, Switzerland, Finland, EU, Spain, Israel, Croatia, Singapore, Australia, New Zealand, Thailand

Science and Technology Research Partnership for Sustainable Development

- India, South Africa, Croatia, Thailand, Indonesia, Tuvalu, Brazil, Egypt, Gabon, Bhutan, Zambia, Tunisia, Burkina Faso, Bangladesh, Ghana, Sudan, Philippines, Vietnam, Peru, Bolivia

http://www.jst.go.jp/EN/CorporationGuide.pdf

Copyright©2003 Center for Research and Development Strategy. All Rights Reserved.

Center for Research and Development Strategy

What is CRDS? Started on 1, July 2003

JST established the Center for Research and Development Strategy (CRDS) in order to fundamentally reinforce the design function of R&D strategies at JST, to enhance our structure as a funding agency, as well as to contribute to R&D strategy designing throughout Japan.

In FY 2006, JST/CRDS has newly established China Research Center (CRC) to enhance mutual understandings between China and Japan.

The missions of CRDS are to:

- 1) Promote dialogues between S&T policymakers and academia,
- 2) Survey S&T fields and draw their "bird's-eye view maps",
- 3) Select important R&D subjects to be funded by the Government, and investigate effective methods for performing R&D on the selected subjects,
- 4) Compare the technology levels of Japan with those of other countries, and
- 5) Propose R&D strategies that can contribute to realization of our vision of future society, enrichment of S&T base, and expansion of research frontier.

9

Copyright©2003 Center for Research and Development Strategy. All Rights Reserved.

Center for Research and Development Strategy

Proposal to S&T Policy Makers

Prime Minister
Council for Science and Technology Policy (CSTP)

Other Ministries (METI, MHLW, MAFF, etc.)
Ministry of Education, Culture, Sports, Science and Technology
MEXT

Strategic Sectors (Generic Goals of Science and Technology)
GREST, ICORP, ERATO, PRESTO

JST/CRDS
Design of R&D Strategy
Direct-General: Hiroyuki Yoshikawa
Principal-Fellow: Ryoji Noyori, Hiroo Imura
Number of Other Fellows*: 42

Fellow: A person who has engaged in research and/or planning of R&D in University, Corporations or other organizations.

10

Copyright©2003 Center for Research and Development Strategy. All Rights Reserved.

Center for Research and Development Strategy

How to make Strategic Proposal?

The vision of our future society
"healthy and comfortable life" "safe and secured society" "learning society" "vital and competitive nation"

Realization of the vision of our future society

Strategic proposals
Strategic Initiative
Strategic Program
Strategic Project

Design and propose effective R&D strategies and recommend them to the Government

Carry out "International Benchmark" investigations

Select important R&D areas in each field (with researchers' community)

Systematically survey R&D fields and draw "bird's-eye view maps" of the fields

ROLL OVER
Researchers' community

11

Copyright©2003 Center for Research and Development Strategy. All Rights Reserved.

Center for Research and Development Strategy

Structure of CRDS

Director-General Hiroyuki Yoshikawa

Executive Director Ken-Kichi Hirose

Principal Fellow (Chair) Ryoji Noyori (President, RIKEN)
Hiroo Imura (Emeritus Professor of Kyoto University, Former Member of CSTP)

Deputy Director-General Tazuo Arimoto, Hidefumi Ueta

China Research Center Director-General Akira Fujishima

Advisory Board

Principal Fellow Kunihiko Nara, Makoto Anashima

Units: Electronics, Information and Communication Unit, Materials Unit, Nanotechnology Unit, Measurement Technologies Unit, Industrial Technologies Unit, Environmental Technologies Unit, Life Sciences Unit, Clinical Research Unit, Policy & System Unit, Overseas Unit, G-To-U

External knowledgeable persons

Research panels

Fellow: A person who has engaged in research and/or planning of R&D in University, Corporations or other organizations.

11

Copyright©2003 Center for Research and Development Strategy. All Rights Reserved.

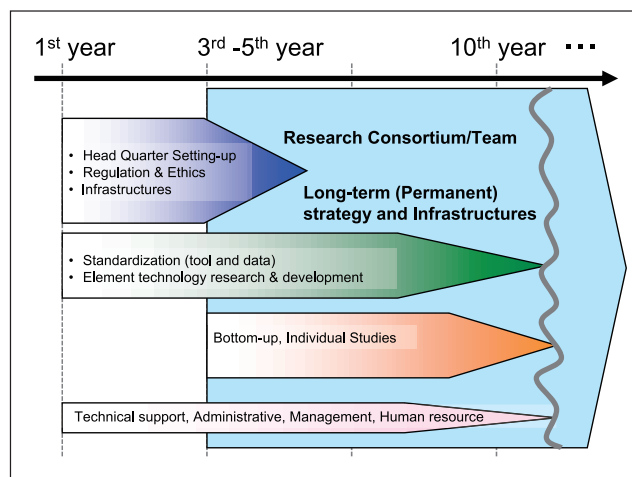
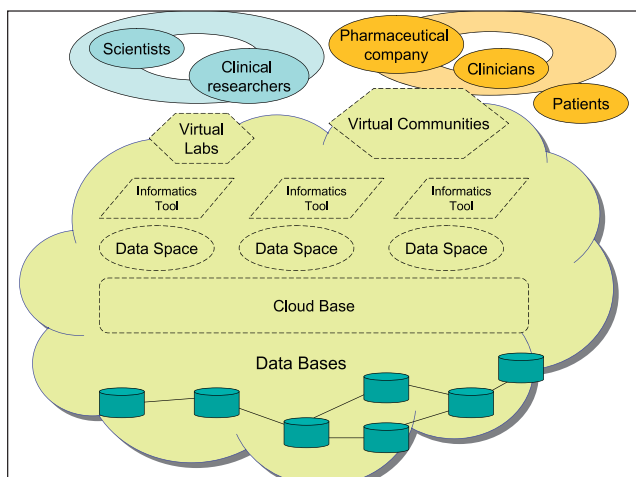
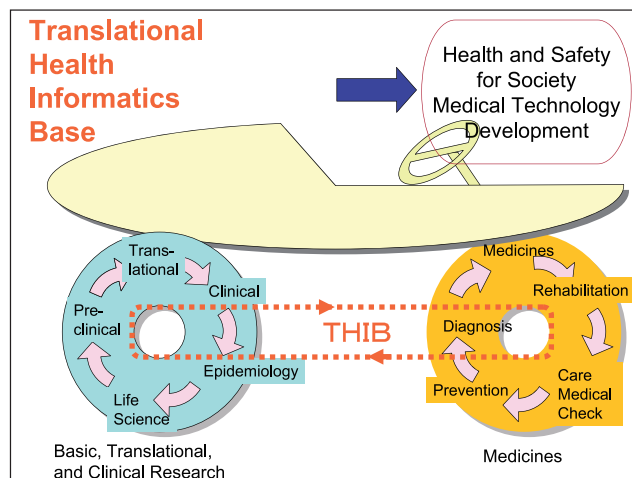
Center for Research and Development Strategy

International benchmark "Translational Healthinformatics Base"

Purpose : Propose a new IT infrastructure to

- establish nation-wide cloud computing system for health science
- facilitate both clinical and research communities to upload and share the biomedical research resource interactively.

Background
The Health Research Promotion Council was established in Japan under the control of the Cabinet Office in 2008. In the strategic plan proposed by the Council, it is recommended to establish "the circulation of knowledge" loop in the medical science and clinical services by clarifying the mechanism of life phenomenon and diseases, converting the clarified mechanism into diagnostic and treatment methods, identifying new research subjects through assessing the effectiveness of the methods, and carrying out research on the identified subjects³



Center for Research and Development Strategy

Bioinformatics in Japan; Projects (1)

"Integrated Database Project" (FY2005-2010)
funded by the Ministry of Education, Culture, Sports, Science and Technology (MEXT)
<http://lifesciencedb.jp/?lng=en>

Subsidiary Program Databases, Life science Databases

The Integrated Database Project

Strategy and Evaluation, Portal Service, Development of Integrated Databases, Integrated Technology, Medical Care Databases

Molecule and Drug, Clinical and Disease, Genome-wide Analysis

Center for Research and Development Strategy

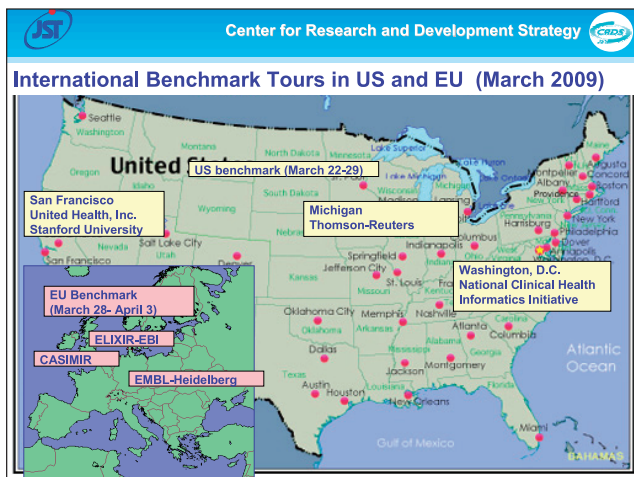
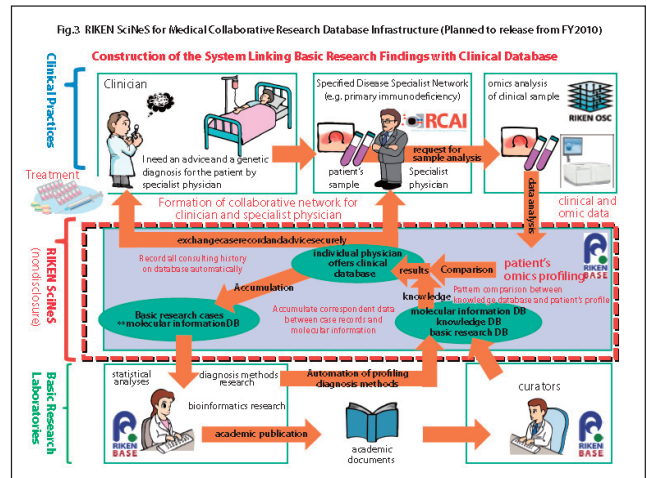
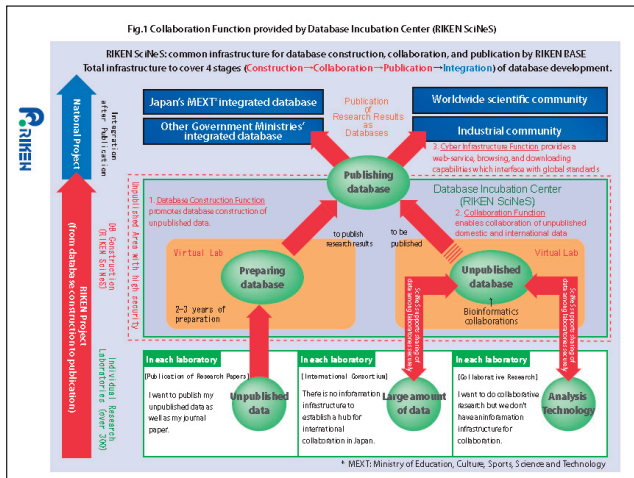
Bioinformatics in Japan; Projects (2)

RIKEN, the Institute of Physical and Chemical Research, Japan
Yokohama Institute, Bioinformatics And Systems Engineering (BASE) Division

The World's First Total Incubation Center for Databases on the Semantic Web

Tetsuro Toyoda
Director

In the field of Life Science, "Data-driven Research", which initially compiles a database of all measured figures of a research sample and makes detailed analysis on individual data later on, has become the mainstream. For this reason, it is essential for each research laboratory to deal with research subjects not only as a 'material', but as a 'database'. Furthermore, supervisors of each laboratory should establish international collaborations with researchers around the world through those large-scale databases. Bioinformatics And Systems Engineering (BASE) Division has constructed a virtual lab center (RIKEN SciNeS) on the Internet with the goal of making such research environment available for every researcher and promote national projects with the aim of integrating Japan's life science databases.



Center for Research and Development Strategy

Questionnaires to EMBL

(to be updated)

- ✓ Current budget
- ✓ Research grant related to technical infrastructure
- ✓ System management
- ✓ System Architecture
- ✓ Incentive for data provision
- ✓ Capability/Flexibility of architecture upgrade (especially for large-scale data base)
- ✓ Human resource (informatics researcher, research assistant, technical staff for maintenance)
- ✓ International collaboration beyond EU
- ✓ Evaluation about the level of informatics research and development in Japan compared with EU
- ✓ Japanese institution, researcher or company well-recognized internationally

22

Copyright©2009 Center for Research and Development Strategy. All Rights Reserved

A2. 世界のバイオ / ヘルスインフォマティクススペースに関する比較調査結果

A2.1. 要素技術研究の国際動向（学術研究）

Thomson Reuter 社の ISI Web of Knowledge を用いて、2010 年 4 月 27 日現在の文献情報から、「bioinformatics」あるいは「clinical decision support」をタイトルまたは抄録に含む学術論文の発表数を 2000 年から 2009 年にかけて調査を行った。これには、若干数の検索目的外の内容を含む文献も入っているが、関連した様々な学術研究成果について、世界におけるおよその実態を示すものと考えられる。その結果をキーワード別に記す。

まず、「bioinformatics」を含む学術論文に関しては、2005 年から 2009 年にかけての学術論文の発表総件数は、739 件（2005 年）から 1114 件（2009 年）と、1.5 倍近くに増加している。このことより、世界的に、当該分野の技術開発ならびに研究の成果が学術的価値を持って拡大していることが示唆される。さらに、論文著者の所属国別（所属が複数国にわたる場合は全ての国をカウントした）に関して分析を行ったところ、米国が突出した発表件数を誇っている他、中国の躍進と英国、ドイツの優位性が確認された（図 A.3.1.1）。特に、中国における発表件数の増加は中国科学技術院による貢献が大きく、研究機関別に見ると、直近 2・3 年に関しては米国ハーバード大学、カリフォルニア大学サンディエゴ校、英国オックスフォード大学等と発表件数のトップ争いを繰り広げている。我が国では国立がんセンターにおける発表件数が顕著であるが、2005 年には世界第 2 位であった論文発表件数が 2006-2008 年にかけて大きく低迷していたことがわかった。

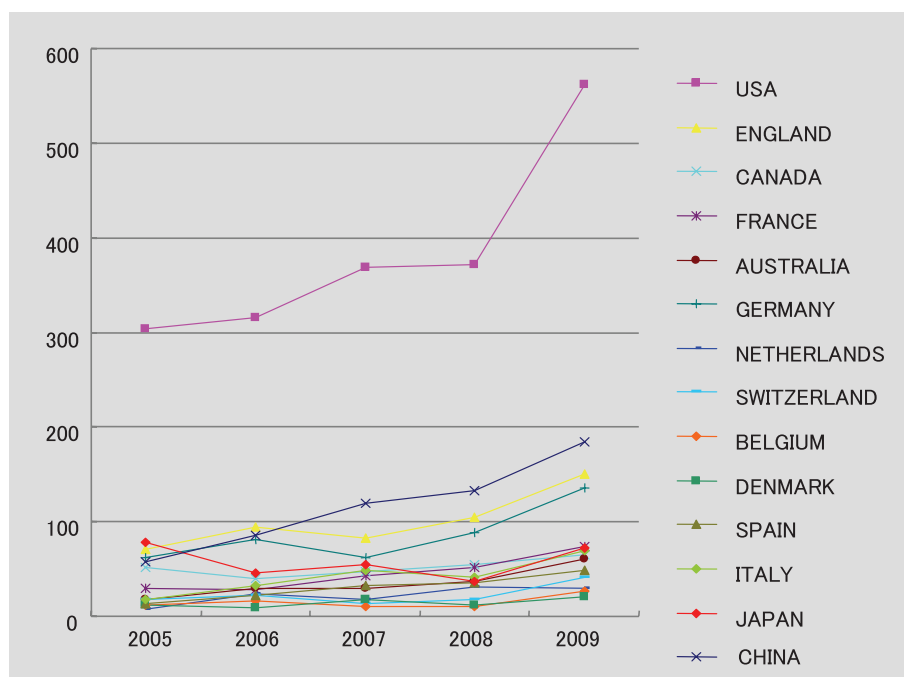


図 A2.1.1. 世界のバイオインフォマティクスに関する学術論文の年次推移(上位 11 カ国)（「bioinformatics」をタイトルまたは抄録に含む学術論文データをもとに CRDS 作成）

また、欧米諸国において、医療情報のインフォマティクスに関する研究成果を反映すると考えられる「clinical decision support」をキーワードに含む学術論文に関しては、学

術論文の総発表件数は「bioinformatics」に比して少ないものの、2005年の217件から2009年の415件と倍増していた。このことより、医療情報を臨床現場で活用する試みは近年盛んになってきていることが示唆される。論文著者の所属国別（所属が複数国にわたる場合は全ての国をカウントした）に関しては、米国が圧倒的に優位に立っている他、英国、カナダが上位を占め、さらにオーストラリアを含めた旧英国連邦諸国において当該分野の研究が盛んなことがわかった。その他、オランダにおいても当該分野における関心が高いことが伺われる（図A.3.1.2）。一方、我が国では2005年に2件の学術論文発表が見られたものの、以降「clinical decision support」をキーワードに含む学術論文が発表されていない。欧米諸国に比べて当該キーワードと異なる用語によって類似した学術研究が実施されている可能性も考えられるが、医療情報に学術的価値を付加するというモチベーション自体が低いことを反映している可能性もあり、今後JST-CRDSの戦略プログラム等によって医療情報を用いた学術研究が推進された際の発表論文数の推移に注意していく必要があるだろう。

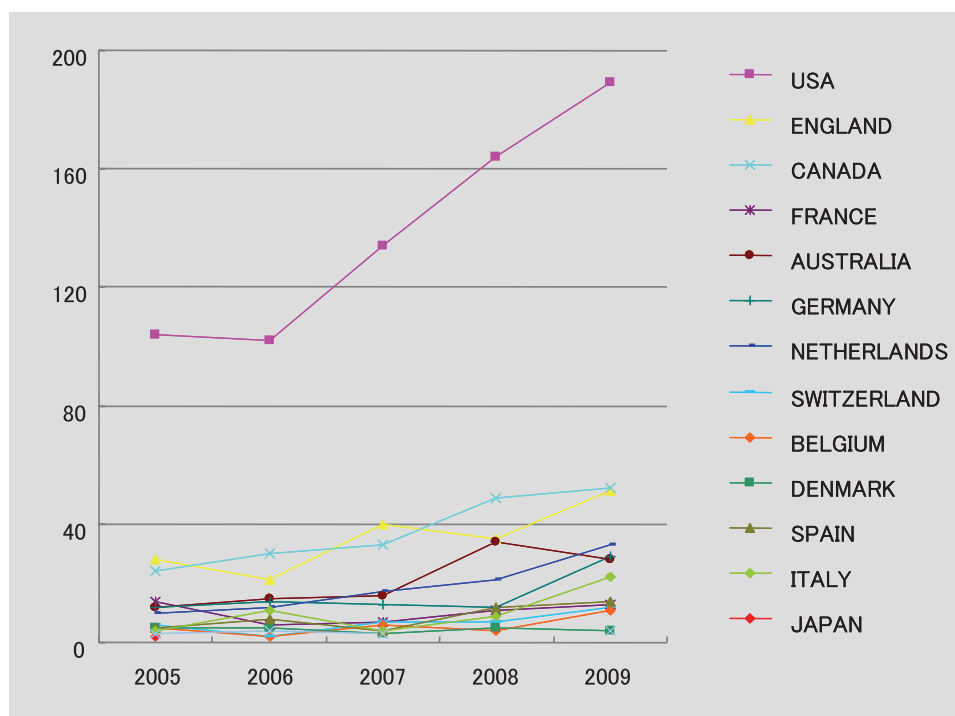


図 A.2.1.2 世界の clinical decision support に関する学術論文の年次推移
(上位 11 カ国 + 日本)

(「clinical decision support」をタイトルまたは抄録に含む学術論文データをもとに CRDS 作成)

A2.2. 要素技術開発の国際動向（推進政策、研究資金）

大規模バイオ/ヘルスインフォマティクスの研究ならびに技術開発に対する公的資金サポートの状況について以下、米国、英国、欧州に関する調査結果を述べる。

<米国>

米国では、国立衛生研究所（National Institutes of Health, NIH）傘下の国立研究資源センター（National Center for Research Resources, NCRR）内において、2001 年から Biomedical Informatics Research Network (BIRN) という研究プログラムが開始されている（http://www.nbirn.net/index_ie6.shtm）。BIRN は全米各地の提携機関の協力を得て医科学研究者向けにソフトウェアツールとインフラリソースを提供している。現在は脳科学に関連した、ヒト脳の MRI 画像と医療情報の連携検索を可能にする Functional Imaging BIRN Tools、マウス脳解剖図のデジタルアトラスを扱う Mouse BIRN Tools、そしてヒト脳の形態計測画像の管理、保管、検索を行う Morphometry BIRN Tools を提供している。Functional Imaging BIRN Tools においては、被験者（患者）各種の MRI 計測画像と医療情報を URL のリンクによって結びつけることによって、アルツハイマー病、躁鬱病、多動性注意欠陥障害、自閉症などの神経疾患をもつ脳の機能的な特徴の抽出、分析が可能になっている。これらのツールを利用するのは脳科学研究従事者に限らず、他の疾患研究に携わる研究者、技術者、医療従事者も多いことがわかっている。

米国に拠点をおくデータベース系企業では、診療データを元に患者の治療フローを分析し、実際の医療プロセスとそのアウトカムを比較することで医療の質の向上、医療経済的な評価に役立てている。また、2009 年の America Recovery and Reinvestment Act（アメリカ再生・再投資法）のなかには Health Information Technology for Economic and Clinical Health (HITECH) Act が含まれ、約 20 億ドルが医療情報のインフラ整備や導入に費やされることが決定している（表 A2.2.1）。この法律によって、健康情報の電子化がなされ、アメリカ国民の健康やアメリカの医療制度を改善することが目標とされている。また、こうした目標の達成には国民の参加や医療従事者による情報の適切な活用が必要とされており、これに向けた技術開発やインセンティブの付与の仕組みづくりが進められている。

米国内でこうした大規模データの構築が進んでいる要因として、

- HIPAA（医療保険の相互運用性と説明責任に関する法律）の施行による医療データの取得や 利活用に関するルールが明示化されている
- 医療保険の大部分が任意加入の民間保険であり、保険加入時の契約書に匿名化した医療データの利用に関する項目が含まれている
- こうした民間保険の業務にとって、大規模データから抽出される情報によって効率的な医療サービスを提供するが経営上も有用である

といった点が挙げられる。また、こうしたデータベースを構築する上で必須の要件として、

- 医師、医療機関、患者（国民）のそれぞれに固有の ID を持たせることが挙げられていた。

一方、これと関連して米国においても課題とされていたのが、データの一意性確保の問題であり、現状の仕組みでは social security number（社会保障番号）をこうしたデータベース内に含めることができないため、転職などで保険者が変わってしまうと追跡できなくな

る、ということが指摘されていた。

また、データは溜めるだけでは活用ができないため、公的な枠組みで大規模データベースを構築するのであれば H-CUP の例のように維持・管理・活用のための中心的な団体を設置し、完全匿名化した上でデータを広く研究で利用できるような仕組みを合わせて構築する必要があると思われる。

表 A2.2.1 米国における医療情報利活用の推進状況

新規プログラム	内容	投資額 (百万ドル)
地域普及センター	医療情報技術の取り扱い等を医療提供者に普及させるための機関を全米で 70 機関設立	643
健康情報交流	地域間での健康情報のやり取りを支援	564
人材育成プログラム	健康情報の専門化の育成など。4 万 5 千人の育成を目標としている。	118
健康情報技術の高度化に向けた研究	情報の機密性、ネットワークプラットフォームの構築など健康情報技術の高度化を目指した研究開発の支援	60
全米健康情報ネットワークなど	健康情報基盤の構築、整備のための投資	64
その他含めた総計		2000

(New England Journal of Medicine 362:382, February 4, 2010 より抜粋)

コラム 米国医療制度の特徴

2010年3月時点で米国の医療は医療保険、医療提供とも民間の占める割合が多い。したがって、医療サービスの費用とそれによるアウトカムに対する説明責任が強く求められる市場性を有しているのが特徴である。一方、医療の特徴として、サービスの直接の受け手と支払い側（医療保険）の分離や、サービスの提供側と受け手間の情報の非対称性があげられるが、こうした特徴をそのままに市場性が持ち込まれることによって「コストの押し付け合い」が起き、医療費の高騰を招いているという批判もある。以下、米国医療制度の特徴を列挙する。

- ・国民の半数以上が就職先企業から補助金を受けて医療保険に加入し、その90%以上が民間保険である。このため、企業にとっても医療費抑制のインセンティブが強い。
- ・国民の約14%、4000万人以上が医療保険を持たない
- ・GDPに占める総医療費の割合は世界最高の16%、額にして約230兆円である
- ・総医療費の約3割が非効率な医療によって失われているコストとされている
- ・4万4000人から9万8000人が毎年医療ミスで死亡しているとされ、うち7000人が投薬ミスによるとされる
- ・ルーチンとして行われている診療内容は、17年前に効果的とされた治療であるとの指摘がされている
- ・オバマ大統領も「医療費の高い分野が必ずしも良い実績をあげていない」ことを指摘している

2010年3月には、オバマ政権下の懸案であった一連の医療改革法案「Patient Protection & Affordable care Act」、「Health Care and Education Reconciliation Act」が成立した。この改革による予想される影響は以下の通り。

- ☐ 3200万人の無保険者が新たに保険に加入し、保険加入率は83%から95%へアップする見通し
- ☐ 保険会社が既往症を理由として保険加入の拒否、重病患者への医療費支払いの打ち切りなどを行うことを禁じる
- ☐ 今後10年間で約9400億ドル（約85兆円）の財源が必要とされるが、高額保険や高所得者に対する増税などで対処

また、2009年2月に成立した「American Recovery and Reinvestment Act」では、医療に関連して以下の取り組みが推進されている。

- ☐ HITECH法で医療情報活用を推進（総額190億ドル）
 - 医療情報基盤の整備
 - 医療情報の利活用のためのインセンティブ付け
- ☐ 医療の効果・効能の比較試験を推進（総額11億ドル）

<欧州>

ESFRI は、European Strategy Forum on Research Infrastructures という、欧州圏における共通の研究課題を発展させるために重要だと思われる基盤技術開発を推進し、それを通して新しい技術革新を誘導する戦略事業のことである。2002 年に事業が開始され、2005 年には、欧州の主な資金配分組織である、欧州委員会の Seventh Framework Program for Research and Technological Development (FP7) において優先的に投資すべき基盤技術の検討が始まり、その結果が 2006 年に発表したロードマップに掲げた目標に反映されている。

(http://ec.europa.eu/research/infrastructures/index_en.cfm?pg=esfri)。本ロードマップでは、Biological and Medical Sciences (BMS) 分野において、今後 10 ～ 20 年の間に重要になると考えられる基盤技術開発について、ワーキンググループの検討をもとに、以下の項目をプロポーザルとして提案した。

(http://ec.europa.eu/research/infrastructures/pdf/esfri/esfri_roadmap/roadmap_2006/bms-report-roadmap-wg-2006_en.pdf#view=fit&pagemode=none)

- European Infrastructure for Chemical Biology
- European Infrastructure for Systems Biology
- Advanced Light Microscopy for Europe
- European Infrastructure for Synthetic Biology
- High Security Laboratories for Emerging and Zoonotic Diseases and Threats to Public Health
- ANAEE: European Infrastructure for the Analysis and Experimentation on Ecosystems

上記ロードマップに基づき、FP7 において以下の 6 事業が開始された。2008 年ロードマップ改訂の際、これらの事業における進捗状況がワーキンググループによって報告されている。

http://ec.europa.eu/research/infrastructures/pdf/esfri/esfri_roadmap/roadmap_2008/bms_report_2008_en.pdf#view=fit&pagemode=none

- Biobanking and Biomolecular Resources Research Infrastructure (BBMRI)
- European Advanced Translational Research Infrastructure in Medicine (EATRIS)
- European Clinical Research Infrastructures Network (ECRIN)
- Upgrade of European Bioinformatics Infrastructure (ELIXIR)
- European Infrastructure for phenotyping and archiving of model mammalian genomes (INFRAFRONTIER)
- Integrated Structural Biology Infrastructure (INSTRUCT)

また、本改訂版においては、潜在的な需要のある研究開発分野として、革新的イメージング技術開発、ケミカルバイオロジーの産業展開、より高い安全管理体制を持つ実験設備

の設立、海洋生物学のさらなる推進、を挙げている。2010年12月にはロードマップ第二版が発表予定であり、それに向けた議論がすでに始まっている。

その他、欧州委員会の科学技術政策に関する情報検索サービス CORDIS (Community Research and Development Information Service) によると、医科学研究、健康科学など、本戦略プログラムに関連した FP7Cooperation (EU 加盟国以外の国の研究者も参加可能な目標達成型国際共同研究) Health 枠において 11 件採択されている。実施期間は 3～5 年、配分額は実施期間が 3 年のものは € 2.8 million 前後、5 年のものは € 10 million 前後である。研究代表者の所属国を見ると英国が 4 件と最も多く、次いでイタリア、スペインが 2 件ずつ、残りは、オランダ、デンマーク、ハンガリーの研究者を代表とするチームに配分されている。EU 各国でも医療情報とライフサイエンス研究の連係を推進している。デンマークでは、病院における患者情報と、分子レベルのオミックスデータを統合する事によって、疾患の病態と遺伝子発現の関係解明を進めている。特に、欧州での基盤整備プロジェクトは、本プロポーザルの THIB と関連しており、国際連携を視野に入れた THIB の構築を目指す必要がある。

付録 参考文献等

- (1) 内閣府 健康研究推進会議 「健康研究推進戦略」 平成 21 年 7 月
- (2) ELIXIR <http://www.elixir-europe.org/page.php>
- (3) 統合データベースプロジェクト <http://lifesciencedb.mext.go.jp/>
- (4) 中村裕輔著 「これからのゲノム医療を知る」 羊土社 2009 年
- (5) 森下真一、阿久津達也編、「バイオデータベースとソフトウェア最前線」 実験科学増刊 Vol.26-No7 2008 羊土社

A3. サイトリポート

A3.1. 米国における関係機関担当者へのインタビュー

○ Agency for Healthcare Research and Quality (AHRQ)

調査日：平成 22 年 3 月 23 日

調査者：興梠 貴英（東京大学 健康医科学創造講座）

山本 雄士（CRDS 臨床医学ユニット フェロー）

調査先概要

- 米国保健福祉省（DHHS）下の機関であり、約 300 名のスタッフが勤務している。医療の質、安全性、効率性、有効性の向上が主なミッションであり、また、医療の公平性も重視している。こうした目的のために医療のアウトカムデータを収集し、エビデンスとして医療界のステークホルダー（患者、医療従事者、支払い側、行政機関など）に提供している。
- 同機関の前身は 1989 年に設立された Agency for Health Care Policy and Research (AHCPR) であり、1999 年に改名、改組され現 AHRQ となった。
- 同機関の予算は約 330mil ドルで、そのうち 80% がリサーチグラントとして外部へ提供されている。これらの外部リサーチにおいて AHRQ はデータ取得の面で支援をしている。したがって、データフォーマットは、基本的に委託先が設定している。
- ライフサイエンス研究戦略との連携はまだ構築されていない。

同機関の取り組み内容

- AHRQ が主導する主なプログラムは以下の通り。ほとんどすべてのデータがウェブ上に公開されている。
 - ・ MEPS (Medical Expenditure Panel Survey) ; 米国における医療の利活用状況のデータベース。トムソン・ロイター社の MarketScan® の一部を用いている。この MarketScan® は NIH と共同で作成されたデータウェアハウスであり、主に 65 歳以下 (Medicare の対象外) の医療情報を集計している。
 - ・ Centers for Education and Research on Therapeutics (CERTs) ; 薬剤による有害事象の削減や安全かつ有効な治療薬の推進を目指す。
 - ・ Consumer Assessment of Healthcare Providers and Systems (CAHPS) ; 医療保険の評価システム。
 - ・ Quality Measures ; 医療の質を評価するための項目を、National Quality Measures Clearinghouse に委託、作成している。
 - ・ Healthcare Cost and Utilization Project (HCUP) ; 主に入院患者へのサービスの質やコストを、米国全体で評価、比較するためのデータベース。1988 年から連邦、州、医療業界が協力して構築している。診療科によって異なる 6 つのデータベースによって構成されている (下記)。同データベースは政府主導で既に 30 年間継続しており、これをもとに退院要約や医療費請求書などの文書が標準化された。ただし、これらのデータは各州で集計されたのち、各州のフォーマットにされうるため、運用上は米国全体で標準化されているともいえないとのコメントがあった。

The Nationwide Inpatient Sample (NIS)

全米 1000 以上の病院から得られた入院データ

The Kids' Inpatient Database (KID)

小児の入院データ

The Nationwide Emergency Department Sample (NEDS)

救急外来受診のデータ

The State Inpatient Databases (SID)

各州で収集している退院サマリーデータを取得している

最新版では 40 州からデータを得ている

The State Ambulatory Surgery Databases (SASD)

各州で収集している外来外科患者データを取得している

最新版では 28 州からデータを得ている

The State Emergency Department Databases (SEDD)

各州で収集している入院には至らなかった救急外来受診データを取得している

最新版では 26 州からデータを得ている

- 各データベースにおいて収集しているデータ項目は 100 以上に及び、たとえば NIS で収集している項目は下記のようなデータを含む。

現在、主要疾患名、二番目に重要な疾患名

主要手技名、二番目に重要な手技名

入院時、退院時の状態

患者背景（性別、年齢、人種、地区別平均年収等）

保険の種類

支払総額

在院日数

病院の種類（所有形態、規模、教育病院かどうか等）

- データ規模については、それぞれ下記となる。

The Nationwide Inpatient Sample (NIS)

2007 年には 40 州、1044 病院から約 800 万件の退院時データを取得している

The Kids' Inpatient Database (KID)

2006 年には 38 州、3739 病院から約 300 万件の退院時データを取得している

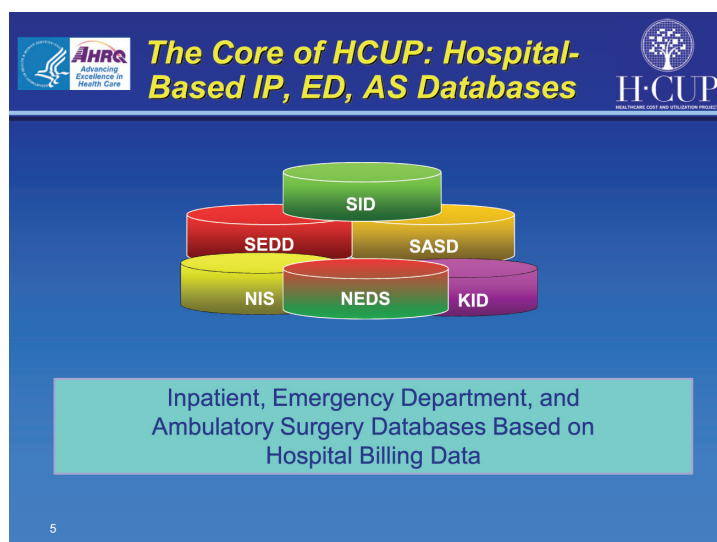
The Nationwide Emergency Department Sample (NEDS)

2007 年には 27 州、970 病院から約 2700 万件の救急外来受診データを取得している

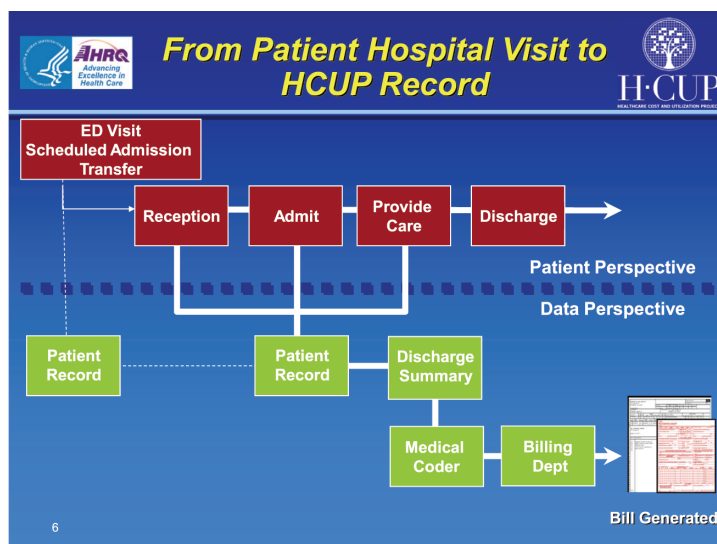
- HCUP に関する情報は、下記のウェブサイト上で詳細に公開され、第三者でもデータ使用に関する同意書に署名したうえで HCUP から購入できる。

<http://www.hcup-us.ahrq.gov/>

- My Own Network, powered by AHRQ (MONAHRQ)；個人向けの医療情報ツール。上記の HCUP を元に作成されている。個人ユーザーが自分の医療情報を入力すると、データの解釈や医療情報サイトの紹介などを行う。
- また、これ以外にも Medicare、Medicaid を統括する CMS の主導で、Hospital Compare というウェブサイトが設けられている。AHRQ もこれに参加している。このサイトのデータは、CMS が各医療機関に必要な情報の提出を要求し取得している。



HCUP に含まれるデータベース



HCUP でのデータ収集のフロー

その他コメント

- 医療情報や質を公開しているとはいえ、それによってどの程度患者の受療行動が変化するのかについてのエビデンスに乏しい。
- 供与されたデータに対価を払うことはないが、個人を対象に診療内容等に関して無料で調査を行うことが難しくなりつつある。

○トムソン・ロイター社

調査日：平成 22 年 3 月 24 日

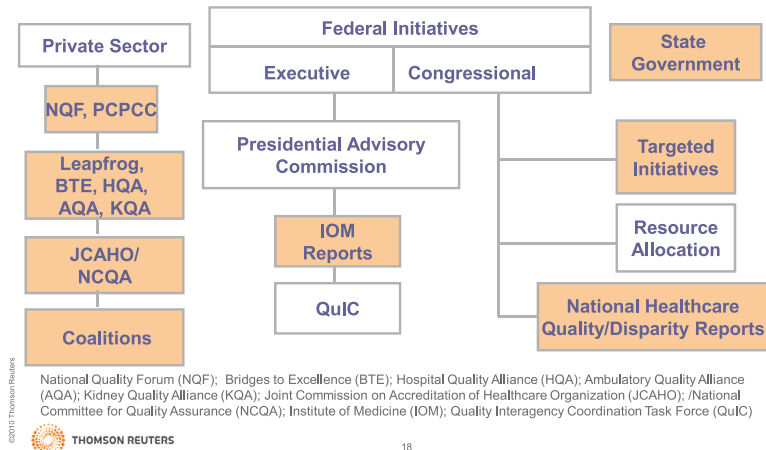
調査者：興梠 貴英（東京大学 健康医科学創造講座）

山本 雄士（CRDS 臨床医学ユニット フェロー）

調査先概要

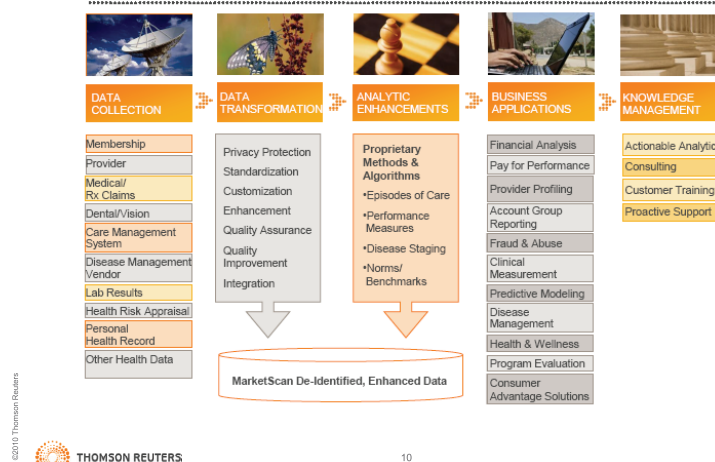
- トムソン・ロイター社は米国ニューヨークに本拠を置く総合情報データベース企業である。現在 93 国に約 5 万人の職員を抱え、年間売り上げは 134 億ドルと同業界第 1 位である。同社は **Healthcare & Science** という部門を持ち、この部門でライフサイエンスから医療まで、広く情報の収集と解析を行っている。
- 医療分野の同社の顧客は、米国政府、30 を超える州政府、保険会社が 100 社以上、病院が 3000 以上、それに大手製薬企業である。同社の持つデータベースには、450 を超えるデータソースから、25 年以上にわたって蓄積してきたデータがあり、これらを各顧客の要望に合わせて編集、解析し、フィードバックしている。
- データの解析などを行う研究チームは約 130 人で、ボストン、ワシントン、サンタ・バーバラ、アナーバーに拠点を持つ。おもな業務は、疾病の重症度およびリスク別調整、健康の面からみた医療の生産性、精神疾患および薬物乱用の調査、国民医療費の動向調査、医療アウトカムの調査、介護の状況調査である。

NATIONAL QUALITY AGENDA: THOMSON REUTERS PRESENCE



同社が米国内で関与する医療の質に関する活動（トムソン・ロイター社資料より）

THOMSON REUTERS CORE BUSINESS MODEL



同社におけるデータ収集から解析、フィードバックまでのフロー
(トムソン・ロイター社資料より)

同社の取り組み内容

- 政府機関である Agency for Health Research and Quality (AHRQ) との連携では、同機関の「病院のコストおよび利用状況に関するプロジェクト (AHRQ' s Hospital Cost & Utilization Project (HCUP))」を受託し、米国の人口の 90% をカバーする医療情報 (コスト、診療内容) を収集、解析してデータを供与している。
- 患者の安全に関わるデータのデータベース構築、臨床判断支援システム (Clinical Decision Support System ; CDS) の構築も行っている。また、後述するように、企業向けに医療保険の費用対効果を公表したり、保険会社向けに病院の質の評価や請求データのチェックを行ったり、病院経営層向けに病院経営の評価をしたりしている。
- CDS とは、医療機関向けのデータベースであり、処方内容や疾病管理方法についてエビデンスに基づく情報 (薬剤の併用状況、退院後の処方内容のチェックなど) を提供したり、医療従事者間の情報共有の支援 (リアルタイムかつ情報の欠損がないなど) を行うことによって、医療提供の質を向上するデータベースである。これは全米の大規模医療機関の 7 割以上が採用している。
- 最近 5 年間で MarketScan for research を用いた研究成果が査読付き雑誌に 200 編以上発表されており、様々な分野における医療評価のための強力なツールとなっている。

QUALITY & COST BENEFITS: CAREFOCUS EXAMPLE

Mortality*	↓	15.6%
Cardiac/Respiratory Arrests*	↓	22%
Med-Surgical to ICU Sepsis Mortality*	↓	38%
CHF Patients Receiving CMS Bundle Measures**	↑	35%
Pharmacy Interventions**	↑	30%
Pharmacy Cost Savings**	↑	\$39,000/month

©2009 Thomson Reuters

*Columbus Regional Nursing IT award ** Baptist Little Rock - Pharmacy Savings

ROI from Baptist was roughly 10:1 in the first year as the system costs in the first year were 100K

THOMSON REUTERS

同社の CDS に基づく改善例（トムソン・ロイター社資料より）

- 基幹データベースの一つである MarketScan® は NIH との共同作業で設計され、1996 年以来延べ 1 億人分を超える医療情報が蓄積されている。データの種別としては、保険者が保有する請求データおよび処方データ、検査値データ、さらに医療提供側（病院など）が保有する経営データなどがある。これらを用いて、医療費の増大が大きな診療科や薬剤、医療行為などの統計を取り、各ステークホルダーと共有している。
- これらのデータ取得は、米国の医療情報保護法にあたる「Health Insurance Portability and Accountability Act (HIPAA)」に基づいて行われている。具体的には、医療機関が診療の際に患者から Informed consent を取得している。取得できない場合は、そもそも医療費請求データを保険者に提供できず、患者が医療費を全額負担することになる。精神疾患などでは、情報共有に一定の制限が加えられる。
- 病院の機能評価として、「100 TOP HOSPITAL」を 1993 年以来毎年発行している。これは、米国の代表的な公的医療保険である Medicare が保有する「Medicare Provider Analysis and Review (MedPAR)」と、Medicare Cost Report を主な情報ソースとし、バランススコアカードに基づいて質、効率性、顧客満足度、財務の 4 つの視点から過去 5 年分のデータを解析し、上位 100 病院を公表するプログラムである。解析のアルゴリズムなどは同社の独自分析に基づいている。以下の 4 つの視点と 9 つの項目を示す。

医療の質

リスク調整後死亡率

リスク調整後合併症発症率

リスク調整後患者安全度（AHRQ による項目）

その他主要項目（医療の質フォーラム（NQF）による項目）

医療の効率性

重症度調整後入院日数

入院費用等

財務

利益率

現金負債比率

顧客満足度

HCAHPS（メディケア・メディケイド・サービス・センター（CMS）による項目）

また、これらのデータをもとに病院のベンチマークを行い、対象病院に対して、医療活動の評価のほか、財務面、労働環境面などからアドバイスを行っている。

- 米国の医療費問題に対して同社では、医療費請求の誤りや不正のチェック、医療機関の運営状況の見直し（たとえば、医師ごとのパフォーマンス評価）、医療サービスの効率的な提供（たとえば患者の疾病管理状況など）の3点から、データ（Medstat®）に基づく解析を行っている。こうした解析とフィードバックによって、受療行動の改善や疾病下院理状況の改善が進んでいる（図参照）。
- 同社はライフサイエンス分野でも、論文や特許、研究開発の状況などに関する大規模データベースを保有している。代表的なものに引用文献データベースである Citation Index や製薬のデータソースである Thomson Pharma などがある。これらと診療情報との連携によるサービス等はない。

同社からのその他コメント

- 医療情報データベースの作成には、医療機関、医師、患者のそれぞれに固有 ID を持たせることが望ましい。しかし米国でもそれは困難で、特に医師の ID は標準化、義務化されていない。
- Medicare の加入者証を貸与して、不正に医療を受給する行為も珍しくなく、こうしたケースを特定するのもいまだ困難である。
- 米国では CDS を、Healthcare Information and Management System Society (HIMSS) や、American Medical Informatics Association (AMIA) が中心となってガイドラインの策定などを行っている。また、政府機関としては AHRQ が資金供与を行っている。

○ユニテッド・ヘルス・グループ

United Healthcare 社

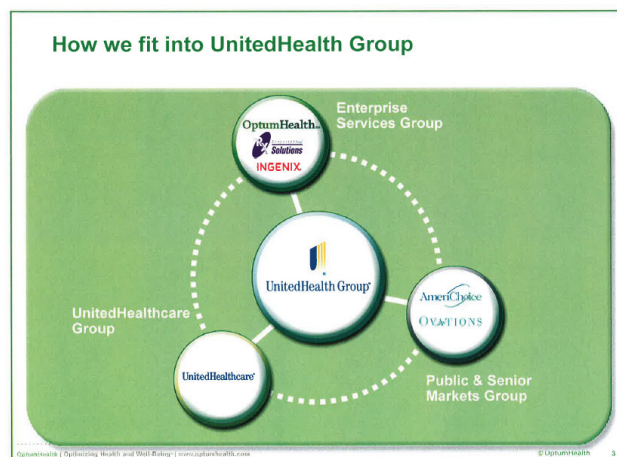
調査日：平成 22 年 3 月 26 日

調査者：興梠 貴英（東京大学 健康医科学創造講座）

山本 雄士（CRDS 臨床医学ユニット フェロー）

調査先概要

- United Health Group は米国の大手医療保険会社グループである。同グループは、保険会社であるユニテッド・ヘルスケア社を中心に、医療情報データベース会社である Ingenix 社、それらを元に保健サービスを提供する OptumHealth 社、製薬企業向けサービスを行う i3 社などを傘下におくほか、5000 を超える医療機関とネットワークを構築する医療コングロマリットである。今回は、大規模医療情報データベースの管理、運用を行う Ingenix 社、その活用を行う OptumHealth 社を訪問した。
- 同社の医療機関ネットワークは全米の 98% をカバーし、およそ 11 兆円分の医療費が同グループの関与の元にある。また、昨年秋から、英国でも英国医療サービス（NHS）との連携によってサービスの展開を始めている。



ユニテッド・ヘルス・グループについて（同グループ資料より）

Ingenix 社

- Ingenix 社は過去 10 年間に延べ 7000 万人分の診療情報を蓄積して解析を行っている。このうち、10 年間連続して追跡できる人数だけでも 1500 万人分である。
- 主な取得データは、請求書データ、処方データ、個人の診療歴等データ（病期、病型などを含む）、検査データである。これらから ImpactPro と呼ばれる同社のデータベースを構築し、個人の疾病リスクおよび予想される医療費、さらには受けるべき医療を解析している。こうしたデータは、保険業のアクチュアリー業界からも認証を受けている。
- 個人の受けた診療内容については、600 を超えるルールに基づき、その妥当性を判断している。こうした実際の診療内容と、エビデンスに基づき推奨される診療内容の差を gap と呼び、その縮小に努めている。
- 複数の医療機関を受診した場合も、同一疾患に対する一連の診療として集計しており、

これによって個人のリスク管理や合併症をふくむ「病態」としての解析が可能となっている。

- 病態ごとにデータを集計、解析することでコスト比較が可能となるほか、コストの上昇要因（コストドライバー）も同定できる。
- Impact Pro の有効例としては、診療ガイドラインの順守率の向上（gap の縮小）、疾病管理の向上、患者満足度の向上などが挙げられている。
- 遺伝子情報は、社会的な受容が未発達との判断から取得、突合していない。
- 2009 年秋より英国 NHS との連携を開始している。英国では、主に病院のデータのみを用いている。これは GP と呼ばれる診療所が、病院と全く異なる、閉じた電子カルテを有しているためである。NHS はデータの提出を要求するものの、フォーマットが標準化されておらず、またデータの適時性も低いため、解析が困難である。

OptumHealth 社

- 同社もユナイテッド・ヘルス・グループに属し、Ingenix 社のデータベースを元に、個人向けに疾病管理、受診勧奨、医療費補助などを行っている。
疾病管理では、特に社会心理学的な要素を考慮し、健康になることへの動機付けを如何に推進するかを研究している。これは、慢性疾患をもつ人の 20 ～ 40% が行動様式に問題を抱えている（たとえばうつなど）というエビデンスと、慢性疾患に合併症が加わった場合に医療費が 2 ～ 3 倍に跳ね上がるという知見に基づいている。
- 具体的には、保険会社や事業主、あるいは個人からの依頼を受けて、該当者の疾病リスクや生活習慣のリスクを解析し、最も費用対効果に優れた介入方法を選択し、サービスを開始する。特に手厚い管理を要する場合には、専属の看護師やケア・マネージャーが用意され、生活習慣の改善（食事・運動）から服薬、さらに適切な医療機関への受診や検査の選択などについての指示を行っている。

○ Stanford 大学

調査日：平成 22 年 3 月 27 日

調査者：興梠 貴英（東京大学 健康医科学創造講座）

山本 雄士（CRDS 臨床医学ユニット フェロー）

医療情報の医学研究下での応用の実例

- すくなくともアルツハイマー病、自閉症、前頭側頭葉型認知症について、死亡患者の脳サンプル、健常者および罹患者の血清が蓄積され、これらの診療情報（発症から死亡前の経緯、合併症など）が同じく蓄積されている。
- これらの情報は Stanford 学内で共有される他、疾患コンソーシアムを共同で組んでいる研究所にも共有可能である。
- ただし、こうした情報はウェブ経由でのオープンアクセスには対応せず、外部からのアクセスには Stanford 大学へ申請をしてもらい、その上で公開するという方式である。
- これらサンプルのバンク、診療情報のバンクは、基本的に寄付金によって賄われている。ただし、有用性が認められた場合には（後述のように）NIH の活動として全国展開されうる。
- アルツハイマー病の診療情報（病歴など）については、UCSF のチームが標準化したフォーマットを用いている。
- この研究成果から、アルツハイマー病の新たなバイオマーカーを開発するベンチャーがスピントアウトした実績がある（このようなスピントアウトの実例として、がんのマイクロアレイがある。がんのマイクロアレイは Stanford 大学で開発された技術であり、当初開発者がリードして寄付金を募って学内にマイクロアレイデータの共有サーバを立ち上げた。後にこれは NIH の所管となり、全米のマイクロアレイデータが NIH に預託されるようになった。しかし、技術の特許は Stanford 大学が有しているため、Stanford 大学にお金が流れる仕組みには変更がなかった）。

A3.2. 欧州における関係機関担当者へのインタビュー

○ European Molecular Biology Laboratory (EMBL) Heidelberg

<http://www.embl.de/index.php>

対応者：Jan Ellenberg (Head of Cell Biology and Biophysics Unit)

Helke Hillebrand (Dean of Graduate Studies)

Lars Hufnagel (Group Leader, Cell Biology and Biophysics)

Phillip Keller (Postdoctoral Fellow, Cell Biology and Biophysics)

François Nédélec (Group Leader, Cell Biology and Biophysics)

Rainer Pepperkok (Head of Advanced Light Microscopy Core Facility)

Patrick Theer (Physics Engineer, Cell Biology and Biophysics)

Stephanie Weil (Online Communications Officer)

調査者：豊田 哲郎（理化学研究所 生命情報基盤研究部門）

堀川 一樹（北海道大学 ニコンイメージングセンター）

富士 珠美（CRDS ライフサイエンスユニット フェロー）

< EMBL の設立経緯、運営資金、ミッション、外部提携（Stephanie Weil） >

- EMBL は、1974 年に設立された欧州共同体の管轄する研究所で、本部をハイデルベルグ市にしている。欧州出身の分子生物学者たちの「米国への頭脳流出」を食い止め、欧州における当該分野の研究基盤を強化することを目的として、当初 10 カ国が参加して設立された。現在は 20 カ国が参加しており、オーストラリアとは連携協定を結んでいる。
- 年間の運営資金は 120 万ユーロで、うち 6 割が参加国からの拠出金、4 割が所属する研究者個人の研究費である。ドイツ政府による年間総出資額は拠出金全体の 3 割ほどで、参加国中最も多い（主な財源はドイツ BMBF）。運営資金は年々増加しており、今後も EMBL の英国支部にあたる European Bioinformatics Institute への英国政府による増資が見込まれている。
- EMBL の主なミッションは「基礎研究の推進（EMBL では、橋渡し研究の前段階になる基礎研究に特化して、ヒトを含めた霊長類を対象とした研究は行っていない、主な実験動物は、ショウジョウバエ、魚類、カエル、マウスなどである）」「新規技術、手法の開発」「加盟国への教育、人材育成をはじめとするサービス提供」「キャリア段階に応じた（研究技術の）トレーニングの提供」「技術移転の円滑化」の 5 つである。
- EMBL の外部提携は CNRS（パリ）、グルノーブル大学（フランス）、バルセロナ、北欧にある。アジアとは特別な提携契約を結んでいる国や組織はない。

（2010 年時点の EMBL 参加国）

ベルギー、クロアチア、デンマーク、フィンランド、フランス、ドイツ、ギリシャ、アイスランド、アイルランド、イスラエル、イタリア、ルクセンブルグ、オランダ、ノルウェー、ポルトガル、スペイン、スイス、英国、（欧州圏以外からの連携国としてオーストラリアが参加している）

< EMBL における研究人材ならびに人材育成 (Stephanie Weil、Helke Hillebrand) >

- 所属する職員数は約 1400 名である。EMBL の人材雇用の特色は、業績、職種等に関わらず全ての職員が同一の職務において最大で 9 年間の契約（3 年契約×3 回更新）しか結べないことにある。この制度が欧州圏の研究人材の流動化の原動力になるほか、EMBL の研究水準、アウトプットの質を高いものに行っているといえる。今までに EMBL 出身者のうち 2000 名以上が欧州圏で大学教授等のパーマネント研究職に就いている他、米国（388 名）、アジアならびにロシア（53 名）、豪州を含む太平洋地区（34 名）、アフリカ（2 名）においても実績がある。
- EMBL 全体で 9 の研究部門、85 の研究室が在籍しており、ハイデルベルグ、ハンブルグ（ドイツ）、グルノーブル（フランス）、モンテロトンド（イタリア）、そしてヒンクストン（英国）の European Bioinformatics Institute、EBI に分散している。そのうち、ハイデルベルグには Cell Biology and Biophysics、Developmental Biology、Genome Biology、Structural and Computational Biology の研究ユニットと所長、副所長直轄の研究ユニットである Director's Research Unit、そして、高額なリサーチツールをシェアする Core Facility が設置されている。その他、アウトリーチ活動や所外研究者、教育関係者への研修などもハイデルベルグの本部が主体的に行っている。
- EMBL の研究施設の利用者は年間 2000 人（外部からの訪問研究も含む）である。技術講習や共同研究を目的に訪れる外部利用者は年間約 1000 人に上り、多くはハイデルベルグ、ハンブルグ、グルノーブルの研究施設を訪問する。
- 2010 年 3 月に Advanced Training Center がオープンした。この施設は主に EMBL が主催するカンファレンス（年間約 50 件）の会場となる会議施設の他、中規模以下の各種会議に対応した会議室や術講習会の受講者のための実習専用実験施設、またその講師やコーディネーターのオフィスを兼ねている。

EMBL では大学院（博士課程）、ポスドク、高等学校の生物学教員、高校生等キャリア別に多様な教育研修の機会を提供している。以下にその概要を記す。

< 大学院（博士課程）教育 >

欧州圏を中心に既存の大学院の博士課程に在籍する学生を Joint Degree Program を締結した 19 カ国の 29 大学（主に欧州圏）から受け入れている。それだけではなく、EMBL が独自に設立している博士課程（3 年 6 カ月）へも学生を受け入れている。現在は 39 カ国より約 180 名が在籍しており、生物学、物理学どちらかの学位を取得できる。EMBL の博士課程への入学試験は年 2 回実施され、受験者の書類審査、受験者の在籍国における基礎生物学分野の評価順位、志望する研究室の定員充足状況の 3 つの要素を総合的に判断して選抜される。合格者が加盟国の学生の場合（例年全合格者の約 45%）は EMBL がフェローシップを授与して学費の面倒を見ている。加盟国以外の学生についてもほぼ全員が外部の奨学金を得て在籍している。教育言語は英語であり（学位論文等は英語で執筆する）、他フランス語、ドイツ語が EMBL の公用語として設定されている。英語を母国語としない学生は EMBL の負担で Onsite Language Course を受講できる。

< EMBL Interdisciplinary Postdocs 制度を通したポストドクキャリア支援 >

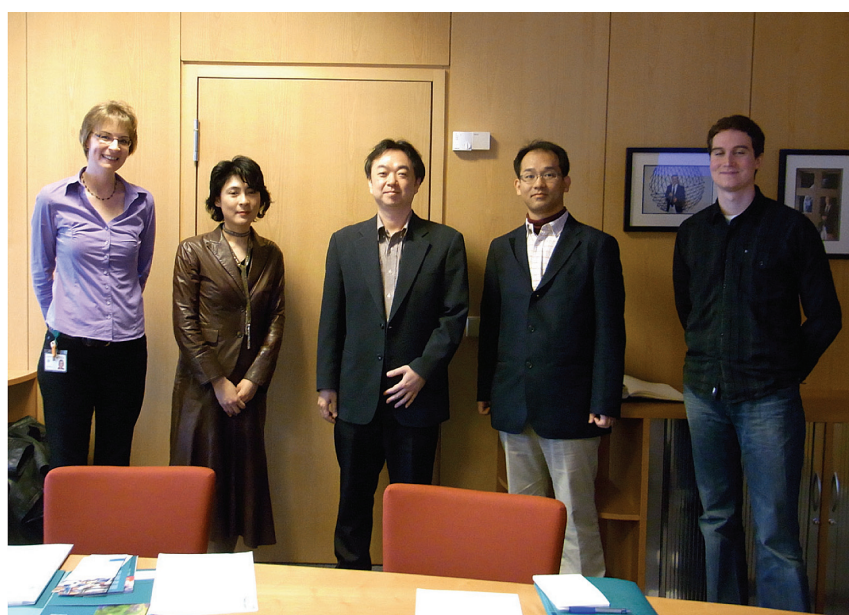
EMBL Interdisciplinary Postdocs (EIPODs) はポストドクが独自の研究テーマによって応募できるフェローシップ制度で、FP7 における Marie Curie Actions と提携しながら、年間少なくとも 20 件以上の研究課題が少なくとも EMBL 内の 2 部門以上に受け入れられるように研究費を助成する仕組みになっている。2010 年現在、EMBL の 5 施設から 200 名が支援を受けている。(http://www.embl.de/training/postdocs/eipod/index.html)

< European Learning Laboratory for the Life Sciences (ELLS) を通した高等学校教員への教育機会 >

ELLS では高等学校の主に生物学の教員が実験施設を使用し、2-3 日の実習コースを受講できる制度がある。講習は一回につき 25 名が定員で、実習にかかる費用は EMBL が負担し、参加者は宿泊費と交通費を負担する。同一教員が再受講することも可能で、年間 5-6 回テーマを変えて実施される。実習は講義と実験からなり、終了後は実験で用いた教材を持ち帰って自身の学校で行うことができることが大きな特徴である。このような試みは主に 11th Grade の生徒を対象とした学校理科教育における教育の質の向上とそれらを通した科学人材の早期啓発を推進する EU における Science in School という取り組みの一環である。Science in School では Quarterly Journal for European Science Teachers という 英文学術雑誌の発行も手掛けている。

< 高校生の生物学研究へのリクルート活動 >

EMBL では他の分野における同様の EU 直轄運営の研究機関と提携して EIRO Forum というものを開催し、高校生への基礎生物学に関する直接的な啓発と教育の機会を設けている。



Stephanie Weil 氏、Dr. Phillip Keller、と記念撮影

(左より Weil 氏、福士フェロー、豊田部門長、堀川准教授、Keller 研究員)

< Cell Biology and Biophysics Laboratory (Phillip Keller, Patrick Theer) >

Ernst H. K. Stelzer 研究室のポスドク研究員と大学院生による対応。Dr. Patrick Theer は 2010 年夏より米国にある民間企業の研究所で Principle Investigator に着任することとなった。

- レーザー光線を物体に水平方向に照射し、さらに上下方向に動かして断層撮影することによって高解像度の立体撮像と分析を可能にする Digital Scanned Light Sheet Fluorescence Microscopy 技術を開発した（参考：Keller and Stelzer, Current Opinion in Neurobiology, 2009）。
- Digital Scanned Light Sheet Fluorescence Microscopy 技術によってメダカの神経系の発達状態やゼブラフィッシュの単一細胞活動などの連続撮像に成功している。
- Stelzer 研究室で行っている発生生物学的アプローチに求められる顕微鏡は「胚全体をカバーできる」「少なくとも 10 ビット以上の情報量を持つ」「1 撮像断面につき少なくとも 107 voxel 以上の撮像速度」「Minimal Bleaching /Toxicity」の仕様を満たさなくてはならない。
- 現在、Cell Tracking（胚の発生を追跡調査する実験）の場合、動物極、植物局側双方からのデータで、3 テラバイト /24 時間にのぼる。撮像された画像データは、地下にあるデータ管理サーバに保管され、そこから専用の「パイプライン」を通して大容量コンピュータによる処理を経て数値化（定量化）しなくてはならない。パイプラインは自動化されており、まず画像が分割され 2 台の CPU による Quality Filters（Automated Image Processing Chart に従ったコントロール画像や False Positive 画像の処理）を経て解析可能になる（参考、Keller et al., Science 2008）。Automated Image Processing のためのプログラムは実験を行う研究者自身が作成し、地下の管理施設に送るようになっている。
- 現在は、一定範囲に分布する細胞群の、発生に伴う移動速度を推定して、細胞を追跡できるシステムを開発中である。また、核の遺伝子発現、細胞の形態、ヒストンチャンネル、細胞膜などに個別の色付けをして構造体の変化を追跡できるようにしたいと考えている。特に、野生群と変異群など、異なる胚の発生過程の比較を行いたい。それには、Atushi' s Marker（理化学研究所 BSI の宮脇敦史副所長の開発した GFP マーカーの総称を指した Keller 氏の表現）を活用して各マーカーの挙動の相関を取って、自動的に数理モデルを構築していくような手法が有効だと考えている。
- High Throughput Experiment Platform for the Virtual Embryo と い う、Double Transgenesis 処理を施した核の遺伝子発現を 24 - 48 時間追跡、再構築したデータプラットフォームを持っている。これらのデータは、全部で 50 テラバイト（1 条件あたり 3 テラバイト）あるが、Dundee University が作成したフリーオンラインソフトである、OMERO Platform を通して画像を取り出し、ダウンロードできるようになっている。OMERO Platform は現存する唯一に近いテラバイトレベルの画像データを共有するためのソフトウェアである。ソフトウェアはオープンアクセスで、付随するメタデータを共有できるソフトウェアもある。Stelzer 研究室では一定期間経過後に撮像データを公開している。

< Cell Biology and Biophysics Laboratory (François Nédélec) >

Nédélec 研究室では染色体の画像解析を行っている。Nédélec 博士は物理学、工学の分野から生物学に取り組むようになった。研究設備の維持を EMBL の研究費で賄い、実際の実験費用は外部資金 (Human Frontier Science Program、Volkswagen 財団、EU FP7 による研究費等) を充てている。

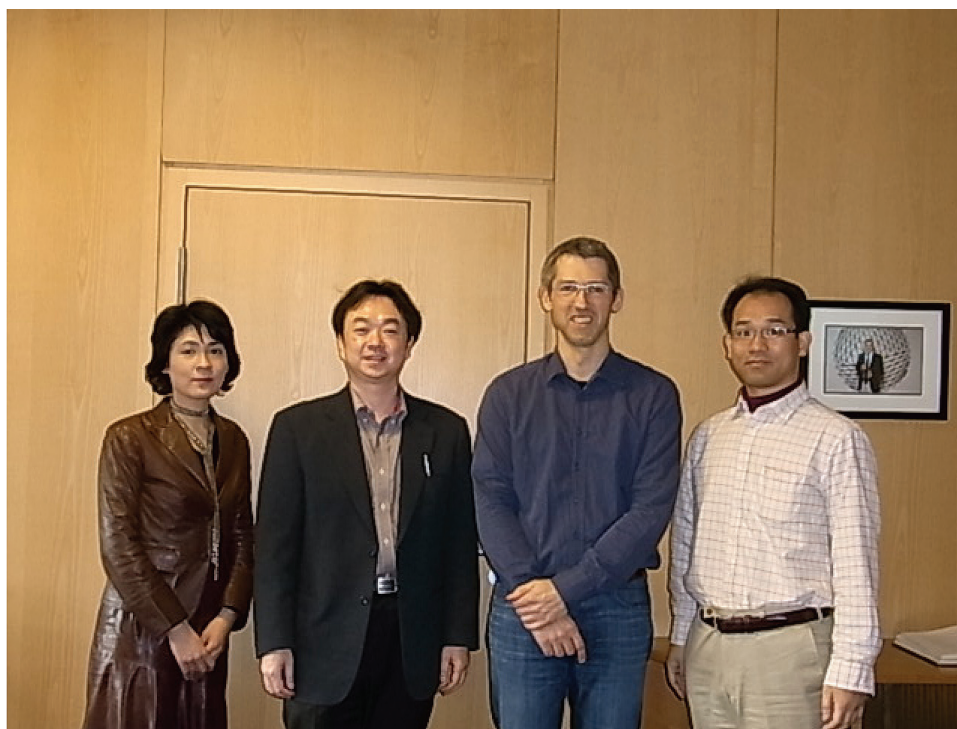
- 微小管 (Microtubule) の画像解析などにもとづいて染色体の挙動に関する数理モデルを構築するには大量のデータを獲得する必要性があり、手動では行うことの困難な解析を自動化して行う必要がある。それらは Matlab など市販のソフトウェアによって作成したプログラムで処理している。
- スナップショット画像と遺伝子シーケンスデータを結び付けるには数理モデルによるシミュレーションが有効であり、そのためのソフトウェア開発がカギを握ると考えている。よりよいモデルの構築には変数の数が重要であり、Nédélec らは15種類の変数を用いてシミュレーションモデルを作ってきた。シミュレーションに用いたプログラムのソースコードは、信頼性が上がったと判断した段階で公開している。Matlab は小規模な実験には適しているが、今後より大規模なシミュレーション実験を行うには不向きな部分もあると個人的には印象を持っている。
- 画像データベースは、それ自体が最終目的、成果物ではない。それらに基づく「解釈 (Interpretation) 重要であり、そのためのツール、ソフトウェアの開発が重要だ。また、画像データの解析において難しいのは品質管理である。
- 画像データに付随するメタデータの共有には OMERO ソフトウェア (前述) の他、オントロジーへの対応にはフリーソフトウェアである Micro Manager を用いている。
- EMBL で実施した研究で得られたタンパクや DNA のデータは EMBL 傘下の組織である European Bioinformatics Institute (EBI) への供出が義務付けられている。しかし、画像データ、シミュレーションデータならびにソースコードは 2010 年 3 月時点で供出が必須ではない。今後は、これらのデータの供出や、ソースコードの継続的な共有、管理のルール作りが必要になってくるのではないかと予想している。

< Jan Ellenberg (Head of Cell Biology and Biophysics Unit) >

2010 年より、EU の FP7CAPACITIES 枠によるデータベースプロジェクトを開始した。自身の研究テーマはヒトの細胞における遺伝子発現機構、特にアポトーシスを起こす遺伝子についてである。

- 2002 年よりヒトの細胞分裂やアポトーシス関連遺伝子の生理機能を解明するために撮影した画像データの解析と標準化に取り組んできた。その手法に関する成果は、2006 年に Nature Method 誌に発表している (Neumann et al., Nature Methods (3) 385-390, 2006)。オリンパス社と共同開発した High-throughput Microscopy は、48 時間で 260000 回分の動画撮影が可能である。
- 大容量の動画データの解析には、バイオインフォマティクスが必要である。Ellenberg らはパターン認識の手法を使って、撮像データから Mitosis の現象を検出し、細胞塊にある核の数をカウントするという手法を用いて、有糸細胞分裂における 2.1 万個のヒト遺伝子機能を調べていった。その成果は 4 月 1 日に公表した (Neumann et al., Nature 464: 721-727.)。論文の公開を待って、これらの撮像データをデータベース MitoCheck を通してオンライン公開した (<http://www.mitocheck.org/>)。データベースは英国にある Wellcome Trust Sanger Institute で作成された。EBI や ELIXIR のデータベースとも統合済みである。今後は細胞内のタンパク間の相互作用を調べていく予定である。
- MitoCheck では、調べたい遺伝子の Phenotype を入力すると、同種の機能を持つ遺伝子群が呈示される。また、コンソーシアムの会員にはデータの共有が許されており、動画データの圧縮版をダウンロードし最解析することや、新たなデータを入力することが可能である。
- MitoCheck 上で公開するデータは、オリジナルなアイデアを持った研究者が、それを公表し終わった後に公開することになっている。第三者がそれを再解析に用いることに特別な契約事項は派生しない。Ellenberg 自身はこの事例を「ELIXIR にとってのパイロット事例」と位置づけている。ある高水準の大きな研究成果を発表したら、そのデータは公表してどんどん使ってもらう方がよい。たとえ競争があってもそれはサイエンスの発展のためであり、それこそが EMBL の理念でもあると考えている。
- 今後の課題は、標準化を担うバイオインフォマティクス技術と技術者の養成である。Microarray のデータと画像データの整合性を図るためのシステムバイオロジーを理解し、技術開発に取り組む人材が重要だ。
- 画像データの標準化に関しては、その手法をシンプルにしないで、皆が使えないことに留意する必要がある。メタデータのオントロジーについても同様で、セマンティックウェブ技術は有効な手段になるのではないか。
- 画像データの適切な管理には、データの圧縮 (Compression) が欠かせないものの、どこかに Raw Data を格納しなくてはならない。データ格納の経費を予算計上できるシステムを導入することは大変重要だ。
- ESFRI 戦略の推進として、FP7 CAPACITIES 枠による EURO-BIOIMAGING のプロジェクトを主宰している (<http://www.eurobioimaging.eu/>)。

- EURO-BIOIMAGING は、医療画像と基礎生物学で得られた計測データ、画像データを結びつける技術の開発を通して、より効率的なデータの管理と処理を実践することを目的としている。23 カ国が参加し、67 の大学、研究所、企業などが公式パートナーとして参加している。参加企業は Zeiss、Siemens に代表されるように機器開発メーカーが多い。現在はデータが分散管理されているが、これを中央で一元管理するためのファシリティへの設備投資である。ELIXIR がシーケンスデータ等従来のバイオインフォマティクス全般の統合とさらなる活用のための基盤整備であるのに対し、EURO-BIOIMAGING は医療画像データを有効活用する目的が明確であるという位置付けの違いがある。
- 現時点で、生物学や医療情報の管理や抽出にクラウド・コンピューティングシステムを使うには、データが大きすぎると考えている。クラウド・コンピューティングは長期的には使いたいと考えているが、その前段階の情報処理のための技術開発があるのではないか。それが、FP7 のライフサイエンス分野における情報通信技術開発への投資に反映されている。



Dr. Jan Ellenberg と記念撮影

(左より 福士フェロー、豊田部門長、Dr. Ellenberg、堀川准教授)

< Advanced Light Microscopy Core Facility (Rainer Pepperkok) >

10 余りの企業の参画による、顕微鏡を用いた技術開発に特化したコアファシリティで、年間 40 名ほどの利用者がある。

- 利用する研究者個々人の要求を満たした仕様のイメージング技術の開発に主眼が置かれている。ソフトウェアも独自に開発する場合がある。現在は、Amira、Huygens、Igor Pro、IPL、ImageJ、Imaris、Metamorph、Microscope Software、ScanR Analysis、VLOCITY などがある。それぞれの企業の製品に独自の特長があるので、一つのファシリティに複数の技術が混在していることに問題はないと考えている。
- 異なる仕様をもつ画像データのメタインフォマティクスについては、ファシリティ内で共通のルールを作り、それに従ったファイルネームを付けることによって単純化して管理している。また、データの中央管理などを担う人材は、コンピュータに関する知識が豊富であるのみならず、オープンな人柄の人材が求められると考えている。

< Cell Biology and Biophysics (Lars Hufnagel) >

高解像度の顕微鏡イメージングによって撮像される画像インフォマティクス研究を展開している。特に、細胞の集合体としての組織が形作られる力学的メカニズムを解明するために、細胞表面上のエネルギーや組織の容量を加味した細胞の力学動態モデルの構築、分析に取り組んでいる。

- 自身が行っている実験の大容量データ（1つの実験につき120の動画ファイル、1データファイルが1.5テラバイト）の格納に苦慮している。現在は、共同研究者扱いとしている民間業者（近郊の町 Karlsruhe に拠点を持つ）の巨大なコンピュータ施設にデータを格納してもらっている（EMBLの公式なデータ管理代行業者の契約は結んでいない）。この業者は15年データを保管したのちオープンアクセスすることになっている。
- （自身が記録したデータのオープンアクセスについて）筆頭著者としての原著論文を最初に出してしまえば、その後にデータを公開することには何の躊躇も自分は感じない。物理学の領域では民間会社が研究者の登録した実験データのオープンアクセスを管理するサービスを行っている（航空機の管制情報をやりとりするOAG Aviationの学術研究版）。このようなサービスに今後はアカデミアも関心をもって取り組んでいくのが重要だと考えている。
- 顕微鏡による撮像データは今後のバイオインフォマティクスの大きな比重を占める。ショウジョウバエの胚の核分裂を例にとると、1回の細胞分裂につき7-8分かかかる現象について、50msecごとに200スライスの断層撮影を行う場合、 $200 \text{ image} \times 0.1 \text{ Hz}$ （撮影頻度が1/100秒） $\times 12 \text{ 時間}$ （全体の撮像時間） $\times 8 \text{ Mb}$ （1imageあたりの容量）で、一晩で約7テラバイトの容量のデータを記録することになる。これらを適正に研究者に提供するための「圧縮」「標準化」は大変重要だ。また「データ記録の長期化」への対応を重要である。現在EMBLハイデルベルグには最大で250テラバイト分の容量を持つデータ貯蔵施設があるが、それではとても足りない。現在でもDATに保管し、ロボットが管理するデータ格納施設がEMBLには残っている。
- 大容量のRawdataから必要なものだけを抽出し、蓄積するソフトウェアは研究グループごとに開発しているのが現状だ。異なるソフトウェアで処理された、同じ胚のデータを再度統合するなどの情報共有ソフトの開発が今後重要になるだろう。直近の例としては、ショウジョウバエのデータ共有を試みているFLYBASE (<http://flybase.org/>)がある。このウェブサイトは研究者コミュニティから自然発生的に立ち上げられ、善意による運営管理の下、フリーアクセスのデータベースとして普及している。それゆえ管理体制には不十分な点も多いが、研究者同士の意見交換によってその改善と維持が図られているという特色も持つ。
- 細胞、組織に対する物理学的なアプローチは世界的にも新しい分野である。自身の研究テーマに関しては、細胞集団の力学動態を正確に定量化するには、細胞集団の表面上のエネルギーと、その容量から数理モデルによって圧力を推定するしか有効な手立てはないという認識を持っている。その点で圧力をValidateするためのマーカー開発は重要である。また、画像データがモデル実験の第一歩を呼

び込んでいることにも留意が必要だ。

- 先にも述べたがデータのシェアは生物学研究の発展上重要である。それと同時にシミュレーションソフトなどのソースコードもアカデミアにおいては共有されるべきだと考えている。EMBL 内の研究者の中には論文公表前でも記録データやソースコードを所内に限って公開する例もある。現時点では、多くのデータベースが特定の目的に沿った小容量のデータに対応したプログラムを稼働させているにすぎないが、それらのアップデート、改良を促進する仕組みが必要だ。Nature Genetics 誌などは、出版社自体が投稿された論文に付随するデータベースを保有、管理することも検討している。学術雑誌を公表したのち、他の研究者がテキストベースでソースコードを書き換えられるような簡便な仕組みも重要である。現在、神経細胞のデータに関するメタ情報の書き込みについて大きな問題は発生していないが、将来的には画像解析の際に役立つメタ情報の標準化などに対応が必要と考えられる。



Dr. Lars Hufnagel と完成したばかりの Advanced Training Center の前で記念撮影
(左より 豊田部門長、Dr. Hufnagel、福士フェロー、堀川准教授)

○ Coordination and Sustainability of International Mouse Informatics Resources (CASIMIR)

<http://www.casimir.org.uk/>

対応者：John Hancock (CASIMIR Partner Project Administrator, Head of Bioinformatics, Mammalian Genomics Unit, Medical Research Council Harwell Institute)

調査者：豊田 哲郎 (理化学研究所 生命情報基盤研究部門)

堀川 一樹 (北海道大学 ニコンイメージングセンター)

富士 珠美 (CRDS ライフサイエンスユニット フェロー)

- CASIMIR は FP6 Life sciences, genomics and biotechnology for health 枠の Meeting Grant による、マウスのデータに特化したバイオインフォマティクス研究者連絡会議開催プロジェクトである。3 年間で 600 万ユーロ助成されるが、会議開催経費（交通費、旅費を含む）と 1 名の事務担当と 1 名のウェブ管理者の人件費が助成対象であり、研究そのものには一切資金が提供されない (Dr. John Hancock 自身の研究費は Medical Research Council (MRC) と Wellcome Trust から得ているとのこと)。このような Meeting Grant は、同業者が集まって連携、協調を図るのに大変有効なものであり、加盟する研究者によるクローズドな会議だけでなく、国際会議の主催などにもファンドされ、有効に使われてきた。
- 2010 年 7 月で FP6 による助成期間が終了し、CASIMIR の名称が使えなくなるが、連絡会議の組織体制と活動は存続させていきたいと考えている。しかし、具体的な財源は決まっていないのが現状である。
- CASIMIR のプロジェクトと FP7 によって開始された ELIXIR のプロジェクト間には現時点で連携関係はない。3 月 30 日に開催された会議（物理学における計算論科学研究の推進へクラウド・コンピューティングがどのような寄与が期待されるか、が主題であった）で初めて関係者と会ったばかりである。(ELIXIR の受け入れ母体である) European Bioinformatics Institute とはデータベース対象がマウスであるか否か、というところで棲み分けをはかってきた。
- データの効率的な収集と解析という観点からクラウド・コンピューティングシステムは有効だと考えている。インフォマティクスにおいては分散型モデルと中央集約型 (Centralized) モデルがあるが、クラウド・コンピューティングは後者に近い。クラウド・コンピューティングシステム開発については国としてはフィンランド、分野としては物理学において進んでいるという印象を持っている。
- 研究者によるデータ提供のモチベーションとしては、プロジェクトの一部としてデータ対象となる研究を取り込むことが重要である。一つの研究が論文として成立するのに必要な実験データというのはどんどん増えている。研究者一人だけでデータを収集、解析するという手法では追いつかなくなってきた。Better for your research by sharing という意識が研究者に広まってくるのはそう遠くないと考えている。

- Medical Research Council Harwell Institute のバイオインフォマティクス部門には 5 名の生物学分野のポスドク、他バイオインフォマティクス (2 名)、物理学 (1 名) のポスドクと 2 名のキュレーター、合計 10 名のスタッフがいる。バイオインフォマティクスを専攻とする学位は英国内ではこの数年に授与されるようになってきた。
- Dr. John Hancock が研究を行なっている Medical Research Council Harwell Institute のデータベース EuroPhenome (<http://www.europenome.org/>) は、現在パンク寸前である。ヴァーチャルな一元管理システム上いかに効率的に情報を格納するか、データの最小化、圧縮技術開発研究の推進が重要だと考えている。所内パイプラインを通して顕微鏡データを処理できるシステムが走り始めたので、その効果をはかっていきたい。自身はデータ提供者の側にたって「いかに提供しやすくするか」を考えた技術開発に取り組んでおり、Ontology によって -ics の情報をヒトの疾患情報に結び付けられないか、開発中である。
- (医療情報と生物学データの連携について) 今後は病理学者などがバイオイメージングデータにアクセスしやすくなるような研究環境の整備が重要になると考えている。バイオインフォマティクスというのは分野間のインターフェースであり、バイオインフォマティクス研究者は仲介者 (Mediator) であるという観点から、データフォーマットの交換、共有は重要である、そのためにパブリックドメインのデータベースを展開していきたいと考えている。
- International Phenotype Database という統合データベースがある。これは大規模データへのアクセス窓口として、登録されているデータベース間の相互リンクをはかっている。このようなデータベースでは技術サポートが欠かせない。
- パブリックドメインによるデータベースの提供や、オープンアクセスには生物学研究成果や蓄積されたデータがバイオテロに悪用される可能性もある。自身のラボでは特別な対策をとっていないが、EBI においてはテロ対策を講じていると伺った。
- 国際間でのデータの収集、共有においては著作権や相手国への説明責任を明確にしつつ、常に協調の構図を維持することが大切である。Data Policy を通して参画機関へのデータ公開を義務付ける、外国向けのリソース提供を有償化する、データ提供者、利用者双方によるデータベースの評価システムなど、体制の整備をきちんと行うことが重要だ。
- 海外の提携機関としては、米国国立衛生研究所 (National Institutes of Health、NIH)、Jackson Laboratory、トロント大学、理化学研究所 (横浜)、などが挙げられる。中国の研究機関との連携も模索している。オーストラリアやシンガポールは医療情報とバイオインフォマティクスの連携が進んでいるという印象を持っている。英国内でも医療情報を統合的に扱うシステムを持つ機関は多い。また、Wellcome Trust Mouse Clinics もある。
- 今後の課題としては、異種間のデータをどのように関連付けるか、遺伝子のフェノタイプと疾患の間にある大量のオントロジーをどのように扱うか、が挙げられる。

(参考) CASIMIR の参画機関

University of Cambridge (英国)

Medical Research Council Mammalian Genetics Unit, Harwell (英国)

Medical Research Council, Human Genetics Unit, Edinburgh (英国)

Helmholtz Zentrum Braunschweig - German Research Centre for Infection Research
(ドイツ)

Helmholtz Zentrum Muenchen - German Research Center for Environmental
Health (ドイツ)

European Molecular Biology Laboratory (EMBL), Monterotondo, Rome X.1 Mouse
Biology Programme (イタリア)

European Bioinformatics Institute (英国)

FLEMING-Biomedical Sciences Research Center Alexander Fleming Institute of
Immunology (ギリシャ)

CNR-Consiglio Nazionale delle Ricerche-Istituto di Biologia Cellulare (CNR-IBC)
Monterotondo Scalo, Rome (イタリア)

○ European Bioinformatics Institute (EBI) / European Life sciences Infrastructure for Biological Information (ELIXIR)

<http://www.elixir-europe.org/page.php?page=home>

対応者： Nicola Slater (Administrator, EBI Data Center)

Gavin Kelman (System administrator, EBI Data Center)

Guy Cochrane (Team leader, European Nucleotide Archive)

Dominic Clark (ELIXIR coordination and participation work package leader/Industry Programme manager)

Graham Cameron (Associate director, EBI/ELIXIR working package 2 leader/Chair of ELIXIR Global Committee)

調査者： 豊田 哲郎（理化学研究所 生命情報基盤研究部門）

堀川 一樹（北海道大学 ニコンイメージングセンター）

福士 珠美（CRDS ライフサイエンスユニット フェロー）

< ELIXIR の設立背景、運営形体、運営資金（Nicola Slater、Graham Cameron） >

- ELIXIR の設立構想は、2001 年にストラスブルグで開催された ESFRI (European Strategy Forum on Research Infrastructures) の準備会議（ESFRI については本報告書 A2.2. 要素技術開発の国際動向（推進政策、研究資金＜欧州＞欄を参照）において持ちあがった。ESFRI では、2005 年に欧州全体の科学技術インフラ整備に重要であると考えられる 35 の ESFRI プロジェクトを同定し、その推進方針を公表した。そのうち、European biomedical research infrastructures (BMS ESFRIs) として、Biomedical 分野におけるインフラ整備が必要なプロジェクトが 6 つ挙げられており、中でも欧州全体を包括する持続的なバイオインフォマティクスコンソーシアムの設立を目的としたプロジェクトが ELIXIR である。
- ELIXIR は、European Grand Challenges というプログラムの中で実施されている。現在の活動資金は、主に EMBL-EBI に FP7 を通して配分された Preparatory Grant から得ており、その他に加盟国からの出資金がある。FP7 からは 2007 年から 2010 年にかけて総額 4.5 百万ユーロが配分される予定だが、現在、Preparatory Grant による活動助成期間の延長について検討が行われている。
- 2010 年 4 月現在の ELIXIR には 13 カ国から 32 の組織が加盟している。Preparatory Grant による助成期間終了後は、加盟国からの出資金のみを財源とする自主運営が必要となるため、各国の研究資金配分機関への出資金の拠出を求めて交渉を続けている。2010 年現在、活動資金を拠出している国はスウェーデン、英国、デンマーク、フィンランドである。ただし、このうち、長期間にわたる資金拠出を決めているのはスウェーデンのみであり、他の国からの明確な長期間の拠出に関する確約はとれていないのが現状である。
- ELIXIR の活動計画は、3 つのフェーズに分かれており、2007 年～2010 年が Preparatory Phase（活動のスクーの決定、パイロット研究の実施）、その後 Construction Phase、Consultation Phase を設定している。現在は参加国へのアンケートなどを実施してその活動ミッションを徐々に確立していく過程にあ

る。ELIXIR のコンソーシアムにおいて EBI がハブ組織となり、参加国の国立組織やその他の研究機関がノード組織として連携するような枠組みを考えている。

- 英国に関しては、国家としての情報セキュリティのリスク管理体制下で ELIXIR を活用していくことが検討されている。国家規模の災害が生じた際にも対応可能な、ネットワーク中心のデータ管理、バックアップ体制の構築を ELIXIR によるインフォマティクス基盤整備を通じて行う構想である。
- Preparatory Phase においては下記 5 つの技術的課題についてのフィージビリティ研究を Working Package13 の活動として実施している。
 - ・ Strategic Review of Cell Phenotype Image Data Resource
 - ・ Pilot of the use of European Supercomputing Facilities
 - ・ Assessment of European Resources for Systems Biology
 - ・ ELIXIR EB-eye Feasibility Study
 - ・ European Genotype Archive
- ELIXIR の活動は（資金調達も含めて）今後のグローバル・インフラ基盤整備戦略の試金石となるものだと考えている。その点から、バイオインフォマティクスの新しいコンセプトの提示や活動資金の集金母体を持つ必要がある。ELIXIR の受け入れ母体である EMBL・EBI は、基礎生物学研究の推進を主目的とする EMBL の組織の中でサービスに特化したミッションを持っている点で特異な組織と言える。ビジネスの面からみても、サイエンスとして扱うことが可能な要素技術を提供するのが EBI の役割だと考えている。

表 A3.2.1. ELIXIR 加盟国による活動資金拠出状況

国 名	資金配分機関名	国内対応機関	拠出額 (百万ユーロ)	期間 (年)
スウェーデン	Swedish Energy Agency Fas Formas the Swedish Research Council VINNOVA	Swedish Bioinformatics Infrastructure for Life Sciences (BILS)	1.7 (19 百万 SEK)	3
英国	BBSRC MRC NERC Wellcome Trust	EMBL-EBI	11.4 (10 百万 ポンド)	特定 せず
デンマーク	Danish Agency for Science, Technology and Innovation (追加出資予定組織 University of Copenhagen, Aarhus University, the University of Southern Denmark, Technical University of Denmark will)	University of Copenhagen Aarhus University the University of Southern Denmark Technical University of Denmark	3.5 (1.5 百万 ユーロの 追加拠出 予定)	特定 せず
フィンランド	Ministry of Education through the Academy of Finland Institute for Molecular Medicine Finland (FIMM) CSC – IT Center for Science National Institute for Health and Welfare (THL)	Institute for Molecular Medicine Finland (FIMM) CSC – IT Center for Science National Institute for Health and Welfare (THL)	1.85 (BBMRI EATRIS への助成 との合計)	2010 年の 1年間

<http://www.elixir-europe.org/page.php?page=home> において 2010 年 4 月 17 日現在までにプレスリリースされた情報に基づき、JST-CRDS が作成

< EBI Data Center におけるデータ管理体制 (Gavin Kelman) >

- ELIXIR の運営母体である EBI のデータセンターには、現在 4.5 peta bytes 分のデータが格納できるデータセンターが設置されている。このセンターは、5 年前に設立された。センターに保管されるデータは原則フリーアクセスであることを前提に管理されている。
- データ管理施設は 9 つにわかれたセクションを WellCome Trust Sanger Institute と共有している。センターには 20 名のエンジニアが在籍しており、ソフトウェア開発や情報通信インフラの整備と維持、改良に取り組んでいる。
- データ管理施設は一日稼働させると、全体で 1.7 Megawatts の電力が必要になる。データ格納スペースには現時点ではまだ余裕がある。

< European Nucleotide Archives (ENA) におけるバイオインフォマティクス (Guy Cochrane) >

- Dr. Guy Cochrane は European Nucleotide Archives プロジェクトにおいて、長期間にわたるシーケンスデータのアーカイブ化を研究している。特に、ここ数年の「情報爆発」に対応したデータの格納課題 (Data versus Storage) の解決を目指している。
- ENA では生データ (シーケンスデータとサンプルデータ)、集合データ (Assembly)、特徴抽出 (Feature Annotation) の段階に分けて、研究者から提供されたデータのブレイクダウン、互換性の向上、局所データの抽出に対応可能なデータ管理を行っている。
- ENA のあたらしいプロジェクトとして、Sequence Read Archive (SRA) という、パブリックドメインを利用したデータの提供、保管、管理の枠組み作りを支援する活動が始まっている。ウェブを介して、データ提供者が大容量の医療データ、個人の実験データなどを提供し、その際公式な契約を結ばないことが特徴である。SRA では、提供されたデータのメタデータの整理と標準化 (Normalization) を行い、データユーザー側へのデータ提供を行う。
- SRA が提供するサービスの利点としては、データ提供者側にはデータの公開に伴う法的手続き等を自身で行う手間が省ける、データの格納場所の節約ができる、などの利点がある (SRA は、データ容量を減らすための Data Compression 技術開発研究費も獲得しており、1 名の専任研究者の他パートタイムの研究者がこれに従事している)。
- ENA では海外の機関として米国 NCBI、ならびに日本の国立遺伝学研究所と提携してきた。SRA では、上記 2 機関に加え中国の研究機関と提携している。国際的なデータの共有提携によって、1 つの実験データセットにつき 6 つのコピーが世界の中に存在することになり、データの持続性が強化できる。地理的に近いところで行うことよりも地球上でデータが分散的に維持管理されることの方がリスク管理的には重要であると考えている。

- 提携機関とのデータのやり取りは様々である。日本側からはまだデータが届く段階にないが、中国の機関はハードディスクにデータを収めて輸送してくる。米国とは、米国内のみで保持するコンフィデンシャルデータを除いたデータを提供してもらう契約を結んだ。
- クラウド・コンピューティングシステムをバイオインフォマティクスに持ち込むことについては、よい試みだと考えている。たとえば、Amazon（書籍、音楽ソフト等のネット流通大手）はダブリンにデータセンターを持っているが、インターネットを用いることで世界的にビジネスを展開している。同様のことは、クラウド・コンピューティングによってバイオインフォマティクスにも適用可能だろうと思う。

< EBI における産学連携の推進戦略（Dominic Clark） >

- EBI では、基礎研究の支援、Advanced training の機会提供、そして技術移転の促進、という 3 つの観点から産業界との連携を図っており、その推進母体が Industry Programme である。現在はアストラゼネカ、ネスレなど 16 企業が Industry Programme に参画している。参画企業とは、共同出資による研究を実施するほか、複数企業から要望のある生物学情報については、中立研究機関として標準化技術を制定し、産業界への浸透を図るなどの活動を行っている。将来的には ELIXIR へも Industry Programme による支援を行いたい。
- Industry Programme では、EBI の施設を利用した Advanced training コース（Industry Workshop）を開催している。Industry Workshop の受講によって、産業界のデータユーザーへ、EBI の事業内容の理解促進や、データベース利用の利点を説明する機会を提供できる。また、アカデミアと産業界におけるデータリソース、データベースの重複なども回避することができると考えている。
- Industry Workshop は 2008 年には 8 回、2009 年には 10 回開催された。2010 年も 10 回の開催が予定されている。ワークショップのテーマは毎回変わり、年 4 回開催される Industry Programme の運営会議に相当する Quarterly Meeting の参加者による話し合いで決められている。その際、産業界からの資金提供などの要因は排除され、あくまでも学術的、産業的に関心が高いと思われるテーマを中立的に選定することが求められている。

(参考) EMBL-EBI の Industry Programme において最近開催された Industry Workshop の テ ー マ (<http://www.ebi.ac.uk/industry/Workshops/previous-workshops.html> より)

- Disease Ontologies and Information
- ChEBI Users Workshop
- Structural biology tools and services provided by the EBI
- The Druggable Genome
- Text Mining and Ontologies
- Computational Identification of Peptides
- Effective use of bioinformatics web services
- Therapeutic Applications of Computational Biology and Chemistry
- micro RNA Informatics
- Systems Biology Symposium
- Alternate Transcript Diversity Symposium
- Semantic Enrichment of Scientific Literature

A4. 国内の状況

A4.1. 国内の関連施策

ライフサイエンスデータの統合化に向けた動きは、第3期科学技術基本計画においても、分野別推進戦略「ライフサイエンス研究全体を支える基礎・基盤課題」の中でも重要なテーマとして位置づけられており、確実に到来する大量データ時代に対応する施策としてすでにいくつか実行されている。主要プロジェクトの1つである、統合データベースプロジェクト(平成18年度～平成22年度;平成21年度政府予算額9億円 平成20年度11億円)は、我が国のライフサイエンス関係のデータベースの利便性の向上を図るため、データベースの統合化及び利活用のための基盤技術開発、人材育成等を行い、データベースの統合的活用システムを構築することを目的とし、現在、(独)情報・システム機構を中心とした関連中核機関が積極的なデータ統合化を進めている。また、バイオインフォマティクス研究の推進(平成21年度政府予算額18億円、平成20年度17億円)が(独)科学技術振興機構で実施されており、解析の効率化・省力化、利用の高度化等を実現するための革新的なゲノム解析ツールの研究開発等、ファンディング機能も活用してバイオインフォマティクス研究を行っている。上記2つのプロジェクトは平成23年度より統合化されることが決まっており、現在その準備段階にある。

公共的なデータ提供ソースとしては、(独)理化学研究所を中心に、バイオリソース事業(平成21年度政府予算額32億円+補正33億円 平成20年度32+5億円)が我が国の中核的拠点として、生物遺伝資源(バイオリソース)の提供を行うとともに、基礎データの蓄積も含めた基盤技術の開発を進めている。データとそれに伴うバイオリソース両者を保有する機関である、ライフサイエンス研究全体の基盤拠点として期待される。

さらに医療分野に特化した大型プロジェクトとして、バイオバンクジャパン計画(個人の遺伝情報に応じた医療実現化プロジェクト)(第1期 2002-2007年度 2002年度は補正、計220億円 第2期 2008年—2012年(予定)平成21年度予算額27億円)が進捗している。通称「オーダーメイド医療実現化プロジェクト」として、2期目に入っており、東京大学医科学研究所に設置されたバイオバンクに収集された患者DNAサンプルを活用し、(独)理化学研究所がSNP(遺伝子の個人差)と薬剤の効果、副作用などの関係や疾患との関係を明らかにする研究等を実施、バイオバンクに収集された疾患について発症に関与する遺伝子や発症メカニズムの解明を目指し疾患関連遺伝子研究の加速を目指している。

医療側における診療情報に関する国の施策としては、具体的に未だ見える形にはなっておらず、内閣官房の高度情報通信ネットワーク社会推進戦略本部(IT戦略本部)でレセプト等の医療情報データベース構築について言及されているに留まっている。実際に診療情報基盤の整備において現状で最も重要な課題は、診療データの標準化を如何に進めていくかであることから、集計する情報項目の決定、その粒度の決定、記述形式の決定などの段階的な整備が必要である。2008年より始まった特定検診制度では、こうした視点が活かされた設計といえ、整備の一例といえる。

インフラとして実働する情報技術研究そのものとして、現在、国立情報学研究所にて運用している、NAREGI(National Research Grid Initiative)上にて、ネットワーク上にスパコン等を繋げ別の計算を別々のスパコンで実行し、それを同時解析する試みがなされ

ており、大量データの大量解析時代に向けた有望な取り組みと言える。

より、対象をフォーカスした、インフラ整備の一例としては、前出の理化学研究所において、理研サイネス（Scientists' Networking System）が、科学者同士をつなぐことに注力し、研究現場のニーズを満たすことを重視して開発されたシステムを構築中である。理研サイネスが無数に提供するバーチャルラボを使うことで、一般利用者は連携研究に必要なデータベースなどの情報システムを自前で立ち上げる必要がなくなる等、コスト面でのメリットも見え始めているほか、バーチャルラボによるデータベース統合化でも着実な成果が得られている。これらのメリットを活かし、「医学サイネス」と呼ぶ基礎研究成果データベースと臨床データベースを連結させる仕組みの構築が平成 22 年度より本格化する。

A4.2. インタビュー対象者から挙げられた国内の主要研究者

インタビュー調査を行った研究者、科学行政担当者から氏名の挙げられた、バイオインフォマティクス分野に関連する日本人研究者は以下の 3 名であった。

中村 春木（大阪大学 蛋白質研究所 附属プロテオミクス総合研究センター 教授）

国際的協力による蛋白質構造データベースの構築、蛋白質および関連する種々の生体分子の構造と物性に関する古典力学と量子化学、および両者を連成させたシミュレーション計算を用いた解明、蛋白質立体構造を基にしたデータベース解析（構造バイオインフォマティクス）、生体分子の生理機能発現の物理化学的理解、X線結晶解析実験による酵素反応機構の解析等を研究テーマとしている。

参考 URL： http://www.protein.osaka-u.ac.jp/rcsfpi/index_pi.html

五條堀 孝（情報・システム研究機構 国立遺伝学研究所 副所長、遺伝情報分析研究室 教授）

ゲノム配列データと遺伝子発現から見た分子進化学、生命情報学に関する研究を行っている。

参考 URL： <http://www.nig.ac.jp/section/gojobori/gojobori-j.html>

宮脇 敦史（理化学研究所 脳科学総合研究センター 先端基盤技術研究コア コア長
ならびに 細胞機能探索技術開発チーム チームリーダー）

新しい蛍光タンパク質を利用した分子スパイの開発、および細胞諸現象の多角的理解、新しい顕微鏡光学系システムの開発、局所的、生理的な Ca^{2+} 動態のイメージング等の研究開発を行っている。

参考 URL： <http://cfds.brain.riken.jp/>

■特定課題ベンチマーク報告書作成メンバー■

福士珠美	フェロー	(ライフサイエンスユニット)
山本雄士	フェロー	(臨床医学ユニット)
勝山光太郎	フェロー	(電子情報通信ユニット)
高野守	フェロー	(ライフサイエンスユニット)
丹羽邦彦	上席フェロー	(電子情報通信ユニット)

※お問い合わせ等は下記ユニットまでお願いします。

CRDS-FY2010-GR-01

特定課題ベンチマーク報告書

トランスレーショナル・ヘルスインフォマティクス・ベース

Benchmarking Report

Translational Health Informatics Base

平成 22 年 7 月 July 2010

独立行政法人科学技術振興機構 研究開発戦略センター 電子情報通信ユニット

Center for Research and Development Strategy

Japan Science and Technology Agency

〒 102-0084 東京都千代田区二番町 3 番地

電 話 03-5214-7484

ファックス 03-5214-7385

<http://crds.jst.go.jp/>

© 2010 JST/CRDS

許可無く複写／複製することを禁じます。

引用を行う際は、必ず出典を記述願います。

No part of this publication may be reproduced, copied, transmitted or translated without written permission.

Application should be sent to crds@jst.go.jp. Any quotations must be appropriately acknowledged.

