

総括チームおよび倫理・規制チームにて取り組む。

(3) 収集データの信頼性、継続性

診療データに関する現状の課題としては、①病名が医療費請求のために便宜的な病名となっている、②病院や医師によって病名の付け方に相違がある、といった課題がある。医療用語をはじめとする標準化が必要であり、その前段階として、医療辞書の整備やオントロジーの整備を加速化する。また、収集したデータがプロジェクトの終了によって保管・維持に必要な予算措置が停止し、散逸してしまうようなことがないように、国としてデータを恒久的に維持する仕組みを整備する。例えば、省庁横断でTHIB運営センター的組織を設けるといったことも考えられる。

総括チームにて取り組む。

(4) インセンティブ設計

データ提供者、病院およびそのデータを利用する研究者、研究機関、保険者、医療・健康産業、行政など多様なステークホルダーのインセンティブ（メリット、デメリットなど）を整理し、課題を見つけ出し、そこを解決するシステム設計をする必要がある。

例えば、どこまでのデータを収集するかという観点から、疾患別のデータベースとした場合には、一定条件を満たした疾患群の患者に関して、インセンティブやコストパフォーマンスの説明をし易いといったことがある。また、収集されたデータに対する受益者（データを利用して研究する研究者等）がいる大学病院、医療機関と、当面はデータを提供するだけの医療機関（個人病院等）の違いを考慮しなければならない。

総括チームにて取り組む。

(5) データベースの著作権、利用規約

データを収集し公開するにあたって、その著作権が課題となる。公開されているデータを安心して利用するためにも著作権上の課題をクリアしておかなければならない。

せっかく収集したデータが、特定目的のための利用に限られる現状を変えていく必要があり、倫理の面も含め議論を深めていく。

総括チームにて取り組む。

(6) データの公開

データを公開することに関しては、共有するデータに対してレベル設定を行い、レベルに応じたアクセス制御の実施などを検討する。米国NIHでは、研究データの公開をグラントの要件とするような取り組みもあり、こうした取り組みに関して検討を進める。

総括チームにて取り組む。

(7) 医師会や学協会との連携

THIBへのデータ提供に関する啓蒙や推進に関して、医師会や学協会での推進を実施する。また、学協会などでのデータの標準化に関する活動に対して、ファンディングを実施し、標準化を強力に推進する。ツール連携・データ標準化チーム、総括チームにて取り組む。

4.3 人材育成

(1) 情報研究者の誘導

インフラ系の構築はもちろんのこと、バイオインフォマティクスやクリニカルインフォマティクス分野での情報関連の研究者が必要とされる場面がますます増えてきており、ファンディングにより研究者を誘導し育成を図っていくシステムづくりが重要である。

(2) 研究支援人材の育成

医師に代わって収集データに関する説明を行い、インフォームドコンセントを得る作業に従事するメディカルコーディネーター、収集したデータを整理しデータベースとして利用しやすいものにしていくアナレーター、データベースの運用管理者など、実際の研究に必要な環境を整備・維持していく人材の育成もあわせて実施する。診療情報管理士といった公的資格に準ずる研究情報管理士といった資格制度も検討すべき課題である。

(3) 人材の継続性

トランスレーショナルヘルスインフォマティクスの人材育成は既存の大学組織の中ではサポートを受けにくく、現状では競争的資金に依存せざるを得ない。境界領域的人材の養成、雇用、活動を支援し、講座に定着させる仕組みが不十分である。プロジェクトの切れ目におきるキャリアパスの中断は大きな問題である。米国の大学では競争的資金が切れても一定期間研究者をPI (Principal Investigator) が雇うことができる資金制度があり、我が国でも検討すべき課題である。

5. 科学技術上の効果

医学ならびにライフサイエンス分野において期待される効果としては、以下がある。

(1) 関連した研究分野間の情報関係がスムーズになることによる新しい知識の発見

大規模なコホート研究や、臨床データの蓄積をベースに総合的研究を実施することで、疾患理解や予防に関する新しい知見が得られることが期待できる。これらは新たな医薬品等の開発の促進、効率化にもつながる。

(2) パーソナルゲノム解読を用いた個別化医療の実現に向けた技術の進展

ゲノム解析技術において、高速シーケンサが開発され、解析チップも従来に比して高性能化がすすんでおり、近い将来には1000ドル程度で個人の全ゲノム解析が可能になる時代がくる。パーソナルゲノム情報にもとづいた個別化医療（オーダーメイド医療）への適用が可能となる。疾病と遺伝子の関連性解析に基づいた薬剤の特定、投与方法などの解明が進展する。

(3) EGM (Evidence Generating Medicine) の推進

情報基盤の構築によって、基礎研究領域での発見、エビデンスをベースとした医療、すなわちEBM (Evidence Based Medicine) を補完することができる。それだけでなく、情報を循環させるという観点から、医療側でのデータ（診療プロセスや結果のデータ）をエビデンスとして生成し、研究に活用するという考え方EGM (Evidence Generating Medicine) をライフサイエンスの現場で実践することも可能となる。

また、情報技術分野において期待される効果としては、

(4) 情報技術フロンティアの開拓

シーケンサのデータ出力の高速・大容量化は、計算機処理の高速化を越えたスピードで進展しており、高速シーケンサが生成する大量のデータをいかに高速に処理するかが、情報処理技術としても大きな課題となり、ストリーミング処理技術などで新しい展開が期待できる。

また、Privacy Preserved Miningや多様なデータの処理検索技術、可視化技術、ヒューマンインタフェース技術などにも新しい展開が期待できる。

6. 社会・経済的效果

(1) 医療の質の向上、医療経済的效果

THIBを整備し、収集する診療データのメタデータを標準化するなどし、大規模にデータを収集することを可能とすることにより、医療プロセスとその効果（アウトカム）の定量評価と比較が可能となり、標準的かつ効果的な医療プロセスの構築への足がかりとすることができる。具体的には、診療ガイドラインの整備等により医師の診療の負担軽減が図れ、医療プロセス全体の底上げが可能となり、医療の質の向上につながる。また、そこで得られる診療結果などのアウトカムデータにより、医療の経済的效果の測定も可能となり、医療の効率化につながることを期待される。

(2) 医療産業競争力・研究力の強化

日本の医療・健康産業の産業基盤は強固とはいいがたく、同産業の基礎体力ともいうべき基礎研究から、その発展に必須である橋渡し研究および臨床研究への連携がよりスムーズになり、実際の創薬や診療における研究効率や生産性が向上することが期待される。さらに、こうした情報基盤によって、規制科学（レギュラトリーサイエンス）などへの適用を見据えた研究が充実し、広く医療産業におけるわが国の競争力の強化につながる。

(3) 疫学の進展により最適な治療や予防が可能

情報基盤により信頼性の高いデータを効率よく集めることが可能となり、データの蓄積によりゲノム疫学といった分野を進展させることにより、最適な治療や予防が可能となる。

7. 時間軸に関する考察

4章でも述べたが、医療データ収集の観点から、プロジェクトの形態として以下のようなものが考えられる。

- 疾病を決め、それに関するデータを複数の病院でまず収集し、データ収集上の課題などをクリアにした上で、できるだけ多くの病院からのデータを収集する仕組みを構築する。
- 非常にまれな疾病に関しては、専門医のネットワークを構築し、個別データを収集する仕組みを構築する。

上記観点から疾病を決め、総括機能を担う拠点組織を決め、スモールプロジェクトを立ち上げ、3年～5年間実施し、インフラの設計と基本機能の構築するとともに、データ標準化や共用ツールに関する成功事例を積み上げる。データ標準化のガイドライン、ツール作成のガイドライン、倫理ガイドラインなどが作成される。あわせて、活動をネットワーク化し拡げていくコンソーシアム（コミュニティ）のような仕組みを構築する。また、THIBをクラウド的運用にもっていくための、仕組み（マネージメント）を構築する。

次に、基本機能として構築されたインフラを利用した数個の個別プロジェクトを3～5年間実施し、インフラの拡張・強化を図るとともに個別課題研究に取り組む。さらにコンソーシアム（コミュニティ）を拡充する。個別課題研究に関しては、拠点にて実施する課題と公募による課題推進の2本立てで実施し、拠点に成果が蓄積されるよう配慮する。

こうした国としての情報資源を恒久的に維持管理していく仕組みとして本提案のクラウドベースによる緩やかな連携整備を位置づけ、各省庁の所管するデータベースを含め国として整備していくものとする。

THIBを広め、世界の注目を集める高度なものとすることによって、日本の情報基盤から世界に発信すると同時に、世界の情報基盤の標準化をリードし、世界に貢献することができる。