

一般口演21

セキュリティとプライバシー保護

2020年11月22日(日) 14:00 ~ 14:49 C会場(コンgresセンター4階・41会議室)

[5-C-2-02] 次世代医療基盤法における ePath等のデータ利活用のあり方の検討

*吉田 真弓¹、山本 隆一¹ (1. (一財)医療情報システム開発センター)

*Mayumi Yoshida¹, Ryuichi Yamamoto¹ (1. (一財)医療情報システム開発センター)

キーワード: Next Generation Medical Infrastructure Law, ePath, Anonymously Processed medical information, Opt-out by notification

背景 次世代医療基盤法（以下、本法）が施行された。目的本法の利点は、認定事業者には各機関から個人情報で診療データの提供が可能なる点、提供後は各機関の管理責任がなくなる点、多施設で名寄せ可能で広範囲多角的な比較が可能なる点があげられる。一方で、患者に対する通知によるオプトアウト実施は医療機関の負荷となる。本法が目指す適切なデータ利活用を促さなければならない。方法我々が参画するAMED事業クリニカルパス標準データモデル（ePath）の開発および利活用では ePathと診療データの統合解析のため、ePathリポジトリデータモデルをXMLで定義し、各パスデータを格納するが、これを本法に対応したSS-MIX2をコモンデータモデルとするDB（FAST-HDJ基盤）への格納検証、倫理的観点での検討、ePath普及の運用検討をし、他施設間や多目的なデータの利活用の在り方を検討した。結果FAST-HDJ基盤は柔軟な形式でデータ格納が可能で、e-Path事業の標準リポジトリデータの格納と匿名加工が可能なる点が検証できた。またDPC様式1,EFファイル、レセデータも格納でき、横断的データの解析に有効だった。本法のスキームの活用には通知によるオプトアウトが必要で、運用上の工夫も必要で今後検討を深めたい。考察ePathはFAST-HDJ基盤で比較的容易にハンドリング可能であった。FAST-HDJ基盤は大量データの高速度処理が可能で、飛躍的なデータ増加にも対応でき様々な分析に提供可能と考えられる。各機関での通知によるオプトアウト運用は継続検討が必要で、新型コロナウイルス感染防止の面で人が手渡しせず設置された状態で患者が自発的に受取る方法も考えられるが、本人への「通知」の意味を成すには、理解促進へのポスター掲示、補助資料の配布は必須で、それ以前に本法にある通り、国の責務として国民の理解の増進を図る広報を進めるべきである。

次世代医療基盤法における ePath 等のデータ利活用のあり方の検討

吉田 真弓^{*1}、山本 隆一^{*1}

^{*1}(一財)医療情報システム開発センター

Study of how to utilize data such as ePath in the Next Generation Medical Infrastructure Law

Mayumi Yoshida^{*1}, Ryuichi Yamamoto^{*1}

^{*1} Medical Information System Development Center

Background: In 2018, the Next-Generation Medical Infrastructure Law was enforced to contribute to research and development in the medical field. **Purpose:** The advantage of this law is that each medical institution can provide medical data to an accredited company without anonymizing the personal identifiable information. After providing medical data, each medical institution loses management responsibility. By integrating and analyzing data from various medical institutions, it is possible to make wide-ranging and diversified comparisons. However, as a disadvantage, the medical institution must carry out the opt-out by notification for its own patients, which is assumed to be a burden on the medical institution. **Methods:** In AMED research project, WG5 are conducting the following research to perform integrated analysis of ePath and clinical data. We store ePath and clinical data from participated medical institutions into the FAST-HDJ database constructed by assuming the Next Generation Medical Infrastructure Law, and anonymize those data according to the research purpose. We examined how to operate "opt-out by notification" at each medical institution. **Results & Conclusion:** We have verified that ePath data can be stored in the FAST-HDJ database and that those data can be anonymized adequately. It was also found that the FAST-HDJ database can process a large amount of data at high speed and can cope with a high increase of collected data. The opt-out by notification needs to be devised by each medical institution, and we need to continue to investigate. However, the government that created the law should carry out public promotion activities enlightenment to the Japanese.

Keywords: Next Generation Medical Infrastructure Law, ePath Opt-out by notification, Anonymously Processed medical information.

1. 背景

改正個人情報保護法では診療情報の殆どが要配慮個人情報となり、これまで多機関の診療情報を集めた上で、広義で公益目的として同意取得なしで二次利用されていた医学研究データベース(以降、医学研究 DB)でも、利用目的を明示した上で本人への同意取得が必要となった。しかし、患者に同意を得た上で、医学研究 DB に診療データを収集する場合でも、収集時点では利用用途が明確でない事が多く、そもそも後ろ向き研究では患者に対する同意取得が著しく困難である。¹⁾

このような状況の下で、医療情報の適切な利活用を可能とし、医療分野の研究開発を推進するため、2018年5月に「医療分野の研究開発に資するための匿名加工医療情報に関する法律」(以降、次世代医療基盤法)が施行された。次世代医療基盤法では、医療機関の診療データを、研究や開発に適切に利用できるような仕組みとして、国が認定を行う認定匿名加工医療情報作成事業者(以降、認定事業者)を設置した。この認定事業者は、各医療機関の診療情報をデータ集積基盤に安全に収集して管理を行い、研究者や創薬、医療関連企業からの利用申請に基づき、認定事業者は必要に応じて多施設の診療情報を名寄せの上、適切に匿名加工等の処理を行った上でデータセットとして利用者に提供を行う。

なお、次世代医療基盤法の大きな利点は、認定事業者には各医療機関から個人情報のままで診療データの提供が可能である点、また、診療データ提供後は、各医療機関には、提供した診療データには管理責任がなくなる点、多機関間で名寄せが可能で広範囲多角的な比較が可能となる点などがあげられる。

医療機関が認定事業者に診療データの提供を行う際には、医療機関は、患者に対する通知によるオプトアウト、つまり次世代医療基盤法に基づき認定事業者に診療データを提供すること、拒否の申出がいつでもできること、受付の方法などを記載した紙面を、看護師または医師等が患者に手渡す等、各医療機関で適した方法で患者に対して通知を実施した上で、拒否がない場合は通知後1カ月以降に認定事業者に対して診療データの提供が可能となる²⁾。

我々が実施した AMED 研究事業「SS-MIX2 を基礎とした大規模診療データの収集と利活用に関する研究」の中で、我々は 2018 年 12 月に 3 か月間ひたちなか総合病院にて、ほぼ全ての受診患者に対して通知を行う、通知によるオプトアウトの実証を行った(実証結果については JCM139, JAMI シンポジウム 2019 にて発表済)³⁾。ここでは、入院患者、外来患者共に拒否率 0.05%程で、著しく拒否率は低かった。この実証では、4 か所の診療受付ブロックにて看護師が患者へ通知を行ったが、配布に係る看護師の件数等が発生した。医療分野の研究開発に資するための匿名加工医療情報に関する基本方針⁴⁾を拡大解釈し、各医療機関の運用に沿った通知方法として再来受付機を利用し、配布をオートメーション化した場合には人件費は不要となるが、機器への改修費、場合によっては機器の追加が必要となる。

2. 目的

我々は、2018 年度より AMED 研究事業「クリニカルバス標準データモデル(以降 ePath)の開発および利活用」(研究代表者: 副島秀久/済生会熊本病院)で分担研究担当(以下、WG5)として参画している。WG5 では多施設間の診療データ

で診療プロセス解析やアウトカム解析を行うため、次世代医療基盤法に基づく認定事業者が運営する診療データ集積基盤に、多施設(研究期間内では実証4医療機関)から収集された ePath データ含む診療データを提供し、研究目的に応じて認定事業者が作成した匿名加工データを使って解析等を行う ePath データの収集・解析基盤の構築や検討を行っている⁵⁾。

なお、WG5 では、MEDIS が設立した(一財)匿名加工医療情報公正利用促進機構(以降 FAST-HDJ)は、現在認定手続き中ではあるが、FAST-HDJ のデータ集積基盤(FAST-HDJ 基盤)は厚労省標準規格である SSMIX2 データを中心に DPC データ、レセデータ等を収集し匿名加工して提供が可能のため、解析のしやすさ、研究事業の終了後の継続性、ePath の普及という面を考慮した上で FAST-HDJ 基盤へデータ提供を行う。このように ePath にとって有益で広範囲多角的なデータ分析の仕組みを整備した上で、ePath 研究事業でデータ収集にあたり実施している全患者へのオプトインの同意取得を、次世代医療基盤法に基づく「通知によるオプトアウト」に替えることで現場の負担を軽くし、結果的に ePath の普及と匿名加工医療情報の適切な利活用の推進に繋げることを目的とする。

3. 方法

ePath 事業の WG5 では、ePath と診療データの統合解析のため、ePath リポジットリデータモデルを XML で定義し、実証4医療機関から ePath 定義書に基づいた ePath データ一式を次世代医療基盤法に対応した SS-MIX2 をコモンデータモデルとするデータベースである FAST-HDJ 基盤へと格納検証を行い、加えて次世代医療基盤法に基づくデータ活用を行うにあたっての課題の検討、次世代医療基盤法の制度に基づいた ePath 普及の運用検討、他施設間の統合データと匿名加工方法など、多目的な分析を行うためのデータ利活用の在り方の検討を行っている。本稿では、ePath データ標準リポジットリモデルの FAST-HDJ 基盤対応・データ格納、施設間の統合データと匿名加工手法、次世代医療基盤法の制度に基づいた ePath 普及の運用検討を行った。

4. 結果

ePath データ一式の FAST-HDJ 基盤への格納検証、次世代医療基盤法の制度に基づいた ePath 普及の運用検討、施設間の統合データと匿名加工方法の検討結果を以下に示す。

4.1 ePath の標準リポジットリデータモデルの FAST 基盤対応・データの格納

FAST-HDJ 基盤は、SS-MIX2 標準化ストレージランザクションログ形式データ、DPC 様式1、EF 統合ファイル、D,H ファイル、電子レセプトを収集格納する機能を元々持っているが、追加で ePath 標準リポジットリデータモデルに対応できるよう基盤の改修をし、加えて D,H ファイルを収集可能とする対応を行った。FAST-HDJ 基盤へのデータ取込については、ePath データ(XML)は大きな問題もなく取込ができ、DPC 様式1、EF 統合ファイル、D、H ファイルは取込に問題がないことの検証ができた。また、SS-MIX2 のデータは本事業では標準化ストレージ形式の CSV で格納を行うこととし、取込の検証が行えた。

4.2 多施設間の統合データと匿名加工の検討

多施設間の統合データと匿名加工の検討については、FAST-HDJ 基盤に取り込んだ4病院の ePath データ、SS-

MIX2 データ等を統合した上で、研究目的、内容に沿った匿名加工を研究期間内に行う予定で、現状では、1機関の SS-MIX2 データ、DPC データを利用して、匿名加工前のデータのリスク評価を行い、匿名加工の手法を検討した上で匿名加工を実施し、匿名加工実施後のリスク評価を行った。

4.2.1 本人特定のリスク評価と匿名加工手法

リスク評価、匿名加工の手法の検討で利用したデータ項目、実施方法等は以下の通り。

- 1) 匿名加工前後の個人特定リスク評価を実施。今回のテストは各項目値の出現頻度が 3 レコード未満を個人特定「リスク有」と定義。
- 2) サンプルデータは1施設の ePath ファイル、DPC 様式 1 ファイル(入院期間 2018/12~2019/3)を利用し、データ件数は 327 レコード。
- 3) 評価項目は①入院日数(ePath ファイルの入院日、退院日より算出) ②入院契機 ICD10 コード(DPC 様式 1 ファイル診断情報 より)とする。
- 4) 匿名加工手法は、削除、一般化(ICD10 コード桁数の調整→ ICD10 コード 3 桁を一般化)とした。
- 5) 出現頻度 3 以下のレコードは本人特定リスク有として削除。

4.2.2 リスク評価と匿名加工実施

評価軸を次の 4 つ(①入院日数②ICD10 コード③入院日数+ICD10 コード④入院日数+ICD10 コード(3 桁化))に分類し、出現頻度 3 以上の有効レコード数をグラフにして、図 1~図 3 に表示した。まず、図1の評価項目①②では、入院日数、ICD10 コードの単体での分布を確認した。入院日数については出現頻度 3 以下で削除されるレコードは 4.3%、ICD10 コードは 9.2%で、殆どが有効レコードと判断できた。図 2 の評価項目③は、①②のデータを併せた場合の分布を試みたところ、出現頻度 3 以上の有効レコードが 74.6%に下がった。④では更に、③で下がった有効レコード残存率を上げるため、ICD10 コードを 3 桁化する、つまり匿名加工手法の一般化で加工を実施し、結果、有効レコード残存率が 82.9%に上がることが確認できた。

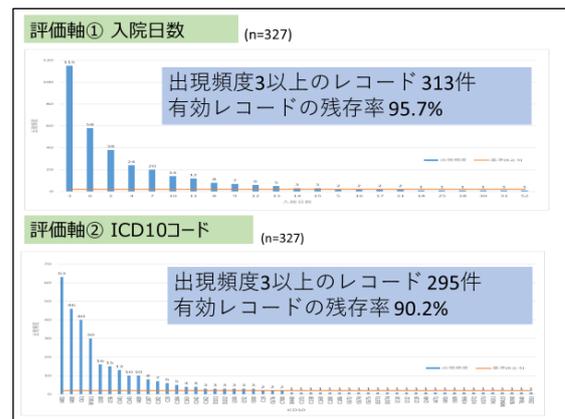


図 1. 評価軸①入院日数 と評価軸②ICD10 コード(単体)

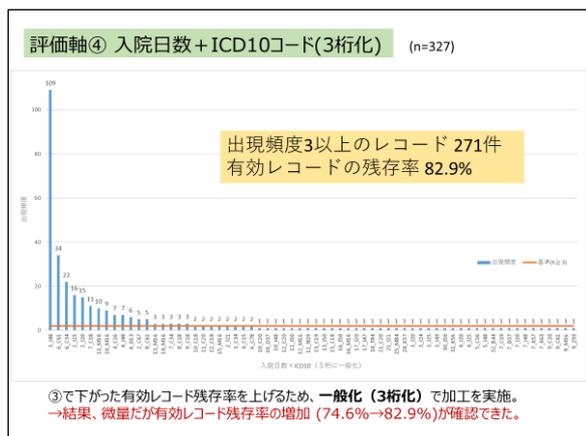


図 2. 評価軸③入院日数+ICD10 コードの有効レコード

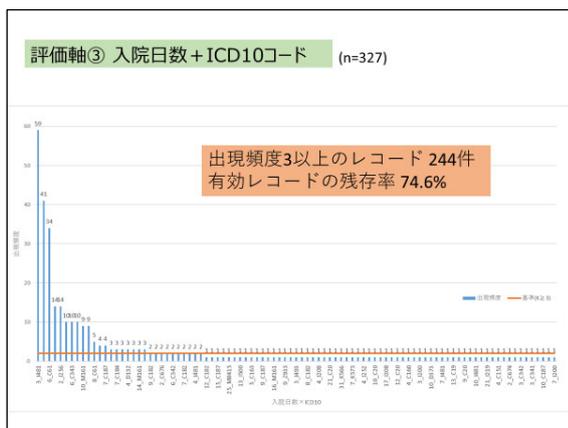


図 3. 評価軸④入院日数+ICD10 コード匿名加工実施後(一般化)の有効レコード

4.3 次世代医療基盤法の制度に基づいた ePath 普及の運用検討

FAST-HDJ 基盤は、収集するデータ形式が標準化され、項目が絞られ揃っているため、診療データを提供する医療機関の負荷が軽く、データを活用する側にとってもデータが扱いやすい。SSMIX-2 標準化ストレージが設置されている医療機関には、FAST-HDJ から提供される D-TransmixCC というアプリケーションサーバを設置し、環境が整えば自動的に FAST-HDJ 基盤へデータ一式が送られる仕組みになっている。しかし、次世代医療基盤法に基づく「通知によるオプトアウト」はいずれの医療機関においても必要であるため、我々が 2018 年 12 月に実施した、ひたちなか総合病院での通知によるオプトアウトの実証で得られた結果(現場の体制整備や運用方法、費用負担、患者の反応やオプトアウト対応)で検討を行い、加えて医療機関の医療情報管理部門の管理者からのヒアリングを行った上で、各医療機関に過度の負担がかからず、法規類の規定に即した実施方法の検討を行った。前提として、以下の 2 点は何れの場合も必要となる結論に至った。

①配布対象を全患者とするか新患、新入院に絞る、もしくは診療科を限定するなど対象を確定させる。②何れの年代の患者にもスムーズな理解を促すような取組を行う。例としては、通知文書以外に、補助資料(仕組みを理解できるような漫画や用語集のパンフレット、解説の冊子など)を配布もしくは院内に設置や掲示などを行う。

上記の実施を前提とした上で、通知の実施には現状のリソ

ースを使う形が望ましく、例えば、新患者には新患窓口で通知し、再来患者に対しては、再来受付の患者が自発的に確実に受取るような場面(受付票やポケベルなど呼出し機器をクリアフォルダに入れる運用)がある場合は、事前に通知文書の中に入れる。月に 1 回実施する保険資格確認の際に実施する。会計窓口で通知を行うなどの運用が考えられ、これらは、認定事業者が支援を行った上で各医療機関に適した運用方法を検討し、無理なく継続できる運用を選択すべきである。

5. 考察

ePath のデータは FAST-HDJ 基盤で比較的容易にハンドリングが可能であった。しかも、FAST-HDJ 基盤で採用している HADB は大量のデータの高速度処理が可能で、将来の飛躍的なデータ増加にも対応ができ、利活用者に対しても効率のよい有益なデータを提供することが可能と考えられる。ePath データだけではなく、画像データも拡張領域に格納が出来ることで、これまで以上に探索的な研究も可能となる。匿名加工手法の検討については、SSMIX-2 と DPC データのリスク評価と匿名加工の検討と実施を行い、一連の流れに問題がないことと、またデータのリスク評価と匿名加工の実施の有効性の評価ができた。多施設で名寄せの上での匿名加工データの分析や、WG4 のデータ解析基盤との匿名加工データの比較検討は今後実施を予定している。

各機関での通知によるオプトアウト運用は継続して検討が必要である。新型コロナ感染防止の面で人が手渡しせず設置された状態で患者が自発的に受取る方法も有効と考えられるが、そもそも「患者への通知によるオプトアウト」の目的は、患者への周知、理解と安心を得ることである。勿論、これに多額の費用がかかることは、医療機関、認定事業者、患者、何れにおいても望ましい事態ではなく、本人への「通知」の意味を成すために、理解促進へのポスター掲示、補助資料の配布は必須で、それ以前に本法にある通り、国の責務として国民の理解の増進を図る広報を進めるべきである。

6. 結語

FAST-HDJ 基盤は柔軟な形式でデータ格納が可能で、ePath 事業の標準リポジトリデータの格納と匿名加工が可能なのが検証できた。また DPC 様式 1, EF ファイル、レセデータも格納でき、横断的データの解析に有効だった。

本法のスキームの活用には通知によるオプトアウトが必要で、運用上の工夫も必要で今後検討を深めたい。

7. 謝辞

本研究を行うにあたり、AMED 研究事業「クリニカルパス標準データモデル(ePath)の開発および利活用」の関係者各位、また日立製作所公共システム事業部公共ソリューション推進第一部各位からの支援に謹んで感謝申し上げます。

参考文献

- 山本隆一, 次世代医療基盤法の紹介と展望, 薬理と治療, 日本臨床試験学会雑誌, 46 号, pp.35-pp39, 2018 年.
- 医療分野の研究開発に資するための匿名加工医療情報に関する法律についてのガイドライン, 2018 年 5 月, 内閣府・文部科学省・厚生労働省・経済産業省
http://www.kantei.go.jp/jp/singi/kenkouryou/jisedai_kiban/pdf/h3005_guideline.pdf ((cited 2020.8.31))
- 吉田真弓, 山本隆一, 次世代医療基盤法に基づく患者への通知によるオプトアウト実施対応の検証報告, 第 23 回日本医療情

報学会春季学術大会,2019年6月,熊本市

- 4) 医療分野の研究開発に資するための匿名加工医療情報に関する基本方針(平成30年4月27日閣議決定),内閣府HP,
<https://www8.cao.go.jp/iryuu/hourei/pdf/houshin.pdf> ((cited 2020.8.31))
- 5) 吉田真弓,山本隆一,「WG5 次世代医療基盤法と臨床パス解析」,第20回日本臨床パス学会学術集会,パネルディスカッション1 臨床パスシステムの標準化の展開,2020年1月18日,熊本市