

る。ただ、かなり医学的に専門性の高いデータリソースであるため、現場の医師などの専門職との綿密な打ち合わせが必要になるであろう。ただ、一部の検査では（LDHやDダイマー値、HbA1c など）、多施設で必ずしも同様の基準を用いていないことがあり、共通データベースを構築する上ではその変換などについても配慮する必要がある。

➤ 薬剤に関連するデータ

薬剤関連のデータの多くはDPCデータにも格納されているものであるが、実際にその薬剤が投与されたかどうかについては一部把握できない部分がある。特に、睡眠薬や下剤、解熱鎮痛剤など、入院患者に対して屯用の形で処方される薬物については、実際の処方について把握することが多くの場合不可能である。屯用薬の処方について知るためには、看護師が行う実施入力が必要となる。今回は、解熱薬の発熱時処方についての二次利用を試みたが、多くの施設で屯用処方の実施についてはその入力信号を把握することができた。また、DPCデータ中の処方データとオーダリングシステム上の処方データとの違いに、どの場所が出されたかどうか、ということについての把握ができるという点がある。たとえば、早期に行わなければならない処方（例：敗血症患者に対する抗菌薬の投与）につい

ては、救急外来で行われた処方か、それとも病棟でオーダーされた処方なのかを把握することで、限界はあるがタイムスタンプの役割を考えると考えられる。

➤ ケア・フロー・データ

入院中のケアフローは、時系列でバイタルサインなどの重要なデータが系統的に格納されているという性質から利用価値の高いデータリソースであると我々は考えた。また、絶対値だけではなく、変動を知ることもケアフローデータの有用な部分である。今回、DPCデータ上の処方データと、同日のケアフローに入力されたデータを組み合わせることで、薬剤がバイタルサインに与える影響についての分析を行うことができた。さらに、「血圧が低い」ということで定義をすることも、「血圧が下がった」という形で定義をすることも可能であり、ここは今後の二次利用を考える上で重要な部分であるといえる。一方、ケア・フロー中に書かれてあることは必ずしも標準化されているわけではないことも今回の研究で問題として指摘されるべきである。呼吸回数はバイタルサインの重要な一部である一方、記載されていないことは少なくなかった。また、異なる手術においては、ケアフローに記載されるデータの種類も異なるため、必ずしも標準化されたデータリソースであるとは言い難い。

➤ 実行に関するデータ

実行に関するデータについては、その信号を把握することができれば、臨床現場での診療プロセスをミクロのレベルで記述することができる貴重なデータとなりうる。しかしながら、実行に関するデータには二つの大きな問題点がある。一つは、その実行がはたして入力されているかどうか、ということであり、もう一つは、その実行がいつ行われたかについて把握できるか、ということである。前者については、たとえば、体位交換が行われたかどうか、聴診が行われたかどうかなどについては記録上把握することは極めて困難である。一方、実行について診療記録への記載が通常行われるような実行も数多く存在する。問題は、実行の証拠を記録する際に多くはテキスト入力での記録となることである。我々は、今回「尿カテーテルの抜去」「安静度の解除」「剃毛」を例に、医療提供の実行を主に診療記録・看護記録をリソースとして二次利用を試みた。この試みは部分的には成功したが、多くの部分で不正確性と不完全性をあらわにしたといえる。定量的なカウントを行う上では、テキストをデータリソースとした二次利用は極めて困難であるというのか今回我々が解析結果を得た上で受けた印象である。

もう一点の課題である実行入力

と実際の実行時間との差については、入力形態を抜本的に見直さない限りは相当困難である、というのが今回の結果から得られる印象である。忙しい臨床現場においては、どうしても実行に関しては事後入力が多くなってしまい、現実との乖離が生じる。今回も、実行についてはほとんど信頼性の高いデータを抜き出すことはできなかった。ここは二次利用を前提とした場合の病院情報の中でも最もハードルの高いものであると考察される。

➤ 退院時サマリなど

一年目の作業が終了した段階で、我々は、入院情報が集約されている退院時サマリからの二次情報の可能性について模索を行った。しかしながら、現時点では退院時サマリに対してあまり大きな期待ができる成果を我々は見出すことはできなかった。まず、検査結果など二次利用のための個別データについては、退院時サマリ用の情報も、データベースから直接抜き出す情報も同様であり、わざわざ退院時サマリを利用する必要はないものである。一方、退院時サマリには、初診時のバイタルサインやプロブレムリスト、重症度評価など、入力時に標準化されていれば二次利用を行う際に極めて有用性の高いデータがまとまっている。問題なのは、退院時サマリ上の情報のほとんどはテキストベースで

入力されており、二次利用に耐えるデータとして格納されていないことにある。退院時サマリの構造上の工夫が二次利用においては必要となる。

➤ 重要だが取得が困難なデータ

病院情報システムとして格納することが困難でありながら、今後医療サービス向上、もしくは臨床研究等を行う上で重要となるデータが数種類が存在する。特に以下のデータについて本稿では考察を行いたい。

◇ 外来診療情報

今回の事業を通じて、外来についてもE Fファイル化できる部分については二次利用が可能であることが明らかとなった。一方、外来においては、入院患者に比較して取得しづらいデータはいくつか存在する。外来データが、入院データに比較してもっとも取得しづらいデータはバイタルサインである。外来データには通常ケアフローが存在しないため、バイタルサインは診療録テキスト上にしか存在しない。さらに、全角・半角それぞれの形で入力されていたり、血圧についても「/」「:」「上××下××」のようにまちまちな形式で入力されているため、二次利用を行うことは極めて困難な状況にある。また、外来においては実施に

ついてもほとんど期待できない状態であった。外来における処置については別途何らかの対策が必要と考えられる。

◇ 病名とケースミックス

二次利用を行う上でもっとも大きなバリアとなるものが主病名である。DPCデータについては資源病名の登録があるため、資源病名を持って主病名とすることは可能である。しかしながら、資源病名が臨床的にも主病名である確率は100%からは遠い位置にある。また、たとえ資源病名が主病名であったとしても、他の病名をもって研究や臨床で特定の疾患グループに加えることができる患者は少なからず存在する。ある患者が特定の疾患を持っていた患者として認識して良いかについては、二次解析だけではなかなか定義することは難しい。

ケース・ミックスについては工夫によって大幅な改善できる可能性が高いが、現時点では問題が大きい。例えば、調整因子として信頼性の高いと認知されている Charlson index のようなケース・ミックスを現在の病院情報から把握することは不可能と考えられる。また、DPCデータについては、様式1で定義されるADLが調整因子としての

役割を持つデータとなりうるが、外来については同様のデータが存在しない。介護保険における要支援度・要介護度などが応用できるかもしれないが、いずれにしても調整因子を病院情報システムから抜き出してくることができるような構造が必要であろう。

◇ 退院後転帰

患者単位での病院情報の二次利用を考えた場合、患者の長期的なアウトカムを把握することは極めて重要である。一方で、長期的なアウトカムについては、既存の病院情報システムを用いて把握することは絶望的であるともいえる。その理由の主たるものは、そもそも病院において長期的に患者をフォローアップすることが困難であること、さらには、生存についての確認に関する記号など、決まった定義のデータ格納方法が存在しないことである。今回我々は、長期アウトカムについて実際には確認しなかったが、これらの重要かつ制限の大きいデータの取得については何らかの手段を考える必要がある。

・ データの連結

➤ IDによる連結

DPCデータを厚生労働省に提出する際に、情報の匿名化が通常行われる。すなわち、病院内で運

用される患者IDは、連結可能匿名化のための統計処理IDに置き換えられる。一方、院内の病院情報システムのデータについては、患者IDが基本的な認識番号となる。今回の研究事業では、すべてのDPC対象の国立病院機構のDPCデータが集約されるNHO本部 総合研究センター診療情報分析部のデータと、各病院の情報システムに格納されている病院情報データを連結させるため、岐阜大学、東京医療センター、および四国がんセンター以外の施設については、院内でのデータ連結は行わず、いったん総合研究センターに情報を集約する形をとっている。すなわち、病院情報システムからのデータを匿名化する際に、患者IDの代わりにDPCデータを提出する際に用いる統計処理IDを病院情報システムのデータに付与するという手順をとる。今回ほとんどの病院においてはその作業を難なく行うことはできたが、一部確認が必要な手順を要した部分があった。DPCデータの提出における連結可能匿名化処理の担当者は、多くの場合病院医事課の職員が担当となることが多いが、臨床研究における個人情報管理の部門との連携が常に可能となるような状況を作ることが望ましいであろう。

➤ 異なるデータリソースにあるデータの連結

- ◇ 今回の参加施設には、異なる病院情報のデータベースを一括するデータウェアハウスの構造を持つ病院と持たない病院があり、ほとんどの病院は後者であった。DPCデータについては、厚生労働省に提出するデータそのものであるため、データの処理に対して職員の負荷がかかることはほとんどなかったが、その他のデータについては、抽出に際して施設に大きな負荷のかかるものと、そうでないものがあった。そして、ほとんどの場合、データベースの構造に詳しいシステムエンジニアの技能を必要とした。一方、岐阜大学のように、すべてのデータ・ベースのデータをソフトウェアを用いて連結させることができるような場合には、必ずしも抽出のたびにシステム・エンジニアに係る必要はなく、ソフトウェアの使用方法を職員のうち数名がマスターすることで、随時多数のデータベースから情報を抽出し、分析することが可能となる。

- ・ 条件定義

- 日時の定義

- 検査部門システムのデータを使用する際に問題となったのは、過去のデータであるために、検査実行の基準をどこに置くかということである。

たとえば、入院時の HbA1c の値を検査部門データから定義して抽出する際に、前向き臨床研究データのように検査実施日が一定していない。その際「入院日±15 日の検査の中で、最も入院日に近い日の検査の値を入院時の HbA1c のデータとする。」というような条件定義を出すことになる。一人一人のカルテを参照しつつ行う後ろ向きの調査においては、一つ一つの個票に対して適切な日付のデータを確認していくことは極めて煩雑な作業であり、正確さにも欠ける可能性がある。今回の試みで明らかになったことは、タイム・スタンプさえ記録されていれば、「最も近い日」や「3つあるうちの中央値」などの条件定義については、十分に定義が可能であり、正確さも手作業で行うよりも優れている可能性が高いということである。その意味でも、入力される際に正確なタイム・スタンプが記録されることが重要である。処置の実施やアウトカムの発生に関する時間定義を入力の時点でいかに行うかが、今後病院情報の二次利用を行う上では重要であろう。

- 対象患者の定義

- 対象患者を定義する上で、薬剤データから定義することを我々は多数行った。なぜなら、特に外来患者に対しては、病名データソースから患者の疾患を特定することが困難だったからである。そのため、“ぜんそく”患者の定義についてはぜんそく治療薬を継続的に使

用している患者を薬剤から定義した。細菌感染症についても、処方された抗菌薬の種類から、疾患や状況について推察を行い形で条件付けを行っている。このような方法で対象患者が定義できたということは前向きに評価することではあるが、同時にこのような方法を用いないと患者を特定することができないということは大きな問題でもある。

さらに、薬の名前を使用して二次利用を行う際に、それぞれの抽出条件に対して薬剤マスタが必要になる。

➤ 病名等の定義

現状の病院情報において最も困難なものは主病名の定義になる。たとえば、「気管支ぜんそく」患者を正確に定義することは難しい。保険病名部分だけで定義することはほぼ不可能であろう。しかしながら、保険病名は病名定義については非常に感度の高いスクリーニングツールであるともいえる。一方、処方される薬剤、たとえば、気管支ぜんそくであれば吸入ステロイド薬などの履歴は、特異度の高いスクリーニングツールであるといえるし、処方期間なども含めて定義すればさらにその特異度を上げることができるであろう。感度が高い条件と特異度の高い条件を組み合わせることによって、サンプルした患者がほぼその病気を持つ患者であるということ特定

することは可能である。ここにおいて、2点残される問題がある。一つは、処方を受けていないが標的となる病気を明らかにもっている患者をサンプリングから見逃すエラーが起こる可能性が高いということである。もう一つは、確かに特定の疾患を持っているが、主病名としての罹患ではなく、対象から外したほうが適切と考えられる患者をサンプリングしてしまうというエラーの可能性である。現時点では、これらのエラーについて具体的な対処法を我々は見出すことはできなかった。これらのエラーを回避する方法として最も効果的と考えられる方法は、入力時に主病名・副病名を正確に記録してもらうことである。ただ、それらを行ったとしても、前向きコホート研究などで設定されるような細かい組み入れ基準を設定して、その基準に準じたサンプリングを行うことは困難であろう。

➤ 文字記号の定義

今回我々は、電子カルテ上の医療文書システムを対象に、そこに格納されているテキストファイルからいかに粒度の高いデータを抽出できるかについての試みをいくつか行った。テキストからのデータ抽出に関しては、データの精度の問題と、個人情報管理に関する問題の二点が考察される。前者については、テキストデータの二次利用は限定された状況において有

用である、というところが我々の結論である。テキストデータからの抽出は、特異度に関しては低いと言わざるを得ないが、感度に関しては、対象を選択すればそれほど悪いものではない。たとえば、「尿道カテーテルの抜去」という事象をテキストから同停止抽出する試みを今回我々は行ったが、「尿道カテーテルの抜去」という事象は様々な言葉のバリエーションがある。「尿道」は「尿」とだけ記載される場合や「膀胱留置」と記載される場合もある。同じように「カテーテル」という言葉も「バルン」という言葉に置き換えられることや、「カテ」と短縮されて呼ばれることもある。これらの言葉のバリエーションを可能な限り“OR”構文で定義することで、見逃しエラーを最小に近づけていくことは可能であるが、ゼロにすることはまず不可能である。今回も、抽出されたテキストから予想外の様々な言い回しが発見され、テキストの条件構文の限界を認識した。また、“OR”構文で条件を緩めていけば、それだけ不要なものがスクリーニングされる頻度も高く、条件定義の特異度は低くなっていく。

テキストファイルからの条件抽出に際して、名詞である「尿道」や「バルン」よりもさらに困難なものが「抜去」のような動詞の定義である。動詞は活用があるためさらにバリエーションが大きくな

るとともに、その行為そのものではなく、行為によって達成された状態を持って記録される（たとえば、「バルン ー」など）こともあるため、抽出条件を設定する側の予測範疇を超えたテキストが出現する可能性が高い。また、漢字やひらがなの組み合わせのバリエーションも動詞の場合は大きいため、感度を高めるためにはかなり条件定義に工夫が必要となる。

テキストファイルを対象に、我々は心房細動患者に対してCHADS2スコアを記録してあるかどうかについてその頻度を調べた。この結果は、実際上の記載があったものを条件定義で同定することが困難であり、感度についても大きな限界を感じた。たとえば、CHADS2スコアを“COH1A1DOS1”などのように、それぞれの頭文字に対して数字を挟んで記載しているようなものについて我々の抽出は見逃していた。診療記録に日常的に書かれるような重症度分類やリスク分類などを、純粋なテキストから正確に抽出するにはあまりにも複雑な条件定義が必要であろう。東京医療センターで使用中の電子カルテシステムにはテンプレート機能があり、文書テンプレートを用いて記録した重症度分類やリスク分類については、抽出時の処理を工夫することでCSV変換が可能であることが分かった。そのため、入力時に必ずテンプレートを

使用するような状況であれば、テキストからの抽出はより正確かつ簡便なものになる。そのためには、実際の日常業務において、テンプレートを使用するほうがテキストを直接書き込むよりも便利であるようなテンプレート運用が必要になるであろう。

問題点の第二点目であるテキストデータの抽出に関する個人情報管理の観点からの工夫について、我々は抽出されたテキストを一括で自動的に匿名化するソフトを開発することで、その要件をクリアしようとした。ソフトは、ほとんどの場合において固有名詞を匿名化することができたが、一方で固有名詞でない名詞や動詞も固有名詞と認識してしまい、そのために必要なテキストに穴が開くこととなった。おそらく、固有名詞同定の条件を緩めることで、誤認識は減ると考えられたが、その選択をとった場合には固有名詞を見逃し、個人情報を匿名化しないままのテキストが残存する恐れがあった。そのため、今回についてはテキストデータを対象としたデータの抽出と解析は東京医療センター内のみで行うこととなった。

・ 個人情報保護の観点からの考察

➤ 単施設における二次利用

個人情報を含むデータを抽出し解析する場合、倫理的な観点からは、そのデータが対象となる施設

の外に出るのか、それとも施設内にとどまっているかということが大きな分岐点となる。もし施設の外に出るのであれば、原則的には連結不可能、もしくは連結可能匿名化処理が必要となる。一方、単施設における二次利用の場合については、かなりフレキシブルな対応が可能である。今回は、多施設データ利用に関する個人情報保護上の観点から、診療記録のテキストを利用する分析に関しては東京医療センターや岐阜大学など、単施設でそれぞれ行い、データを施設外に出さないこととした。匿名化ソフトの限界も含め、やはり単体で個人情報を含み得るデータに関しては、自動的な匿名化を行うことは技術的に大きな限界がある。

単施設での調査の場合にアドバンテージとなりうるものは、直接カルテを閲覧して得られる情報を利用できることである。今回の分析において我々はこの試みは行っていないが、メインのデータ抽出はデータベースから直接抽出し、直接のカルテ閲覧が必要な情報に特化してカルテを閲覧することで、かなりデータをまとめる上での労力は軽減されるであろう。

➤ 多施設における二次利用

多施設における二次利用については、データベースからの直接抽出が可能でデータに関しては、十分に個人情報保護の観点からも運用可能であることが今回の調査

を通して明らかになった。特に、国立病院機構のようなスケールメリットのある法人については、多施設でのデータ収集を行うことで、研究を目的としたものにも、医療サービス向上を目的としたものについても非常に有用なデータを抽出することが可能となる。国立病院機構については、今回対象となった病院群の多くはある特定のベンダーのシステムを利用していたが、どのベンダーの情報システムを採用するかについては病院の裁量に任されているため、必ずしも汎用性が高いとは言えないことを実感した。さらには、同じベンダーについても病院ごとにカスタマイズがなされており、これがかえって標準的なデータ抽出の方法立案に対して障害になることもあった。今後は、ベンダーの規格統一とともに、一定のテンプレートの使用など、入力方式にも工夫が必要であろう。

<各論：目的別の二次利用>

- ・ 臨床評価指標など、医療の質の継続的改善への応用

今後、病院情報の二次利用をより活発に行っていくうえで、その目的としては大きく二点、すなわち、医療サービスに直接資するデータを引き出すことを目的としたものと、研究を目的としたものに分けることができる。今回の我々の事業では、前者および後者両方を対象に分析を行ったが、ほとんどは臨床評価指標

の算出を中心に行った。なぜならば、臨床評価指標の算出は、医療の内容の詳細に踏み込んだきわめてミクロなデータへのアクセスが必要になるために、各施設においてそのデータを高い妥当性を持って生成するためには多大な労力がかかる作業だからである。さらには、その作業は毎年、場合によっては毎月の単位で行わなければならない作業であり、このような作業に対して担当者が支払う負担は絶大なものとなる。データベースから直接データを抽出し解析する方法が適用できれば、これらの負担は一気に解消するため、日常の業務負担を軽減する上では最善の方法であろう。特に、多施設でこれらのデータをそれぞれ抽出するような場合においては、各施設の人的な資源に大きく差があるため、手作業でのデータ抽出を一律に強いることは現実的ではない。データベースへ直接アクセスし抽出を行う場合のメリットは、そのような手間を最小限にするとともに、いったんロジックを作ってしまうと、ほとんど手間をかけずに計時的な分析を繰り返すことができる点にある。今回の事業においても、パイロット分析の段階を経て本解析に至ったが、ロジックの変更を大きく必要としたもの以外は、大きな負担をかけずにデータを抽出することができた。今後、臨床評価指標や病院の医療成績に関する計時的なデータの算出については、DPC データ以外のデータリソースも含めた二次利用の方略を積極的に立案することが推奨される。一方では、これらのデータを多施設間比較する際に障害となりうるインフ

ラの問題については、設計の段階から検討する必要があるであろう。とりわけ、アウトカム変数の入力時点からの標準化については、初期設計なしには把握することが難しいため、どのようなアウトカムを測定し、公開するべきかについての議論を経たのち、アウトカムの記録が妥当性の高いような状態で残るような入力の方法が必要となる。

・ 臨床研究、疫学研究への応用

今回、我々は2つの後ろ向きコホートデザインをプロジェクトの中で試してみた。ひとつは、胃潰瘍に関する除菌プロセスを説明変数とし、再発症状をアウトカムとしたものであり、もう一つは解熱鎮痛薬の処方の種類を説明変数とし、バイタルサインの変化をアウトカムとしたものである。これらの解析については、結果的に限界はあるものの満足いくものが得られた。今回の調査結果からは、臨床研究を目的とした場合においても、少なくとも後ろ向きコホート研究のデザインで行う場合では、十分に二次利用での解析が可能であった。今回は比較的小規模な多施設比較ではあったが、より施設の数を広げることにより、大規模研究を行うことも可能である。

一方では、いくつかの問題点も浮き彫りになった。第一には、前段と重複する問題点だが、アウトカム測定については、原稿の病院情報システムでは大きな限界があるということである。退院後の死亡や再入院、ADLなどについて、共通の基盤を持つデータの収集があらかじめなされていれば、より解析の幅は広がるであろう。しかしながら、研究を目的

としたデータを収集するに当たっては、倫理的にも問題になることが多い。医療サービス利用者との合意の上でこれらのデータについての収集はおこなわれるべきであろう。また、研究事業の場合には多くの場合単体、もしくは短期のデータ収集にとどまる場合が多い。そのため、診療記録の直接閲覧も方法として含めながらデータの収集計画を立案することが大切であろう。

第二には、症例対照研究のように、サンプリングに対して工夫が必要となるようなデータの収集については、まだまだ限界が大きいと考えられる。コホート研究については、サンプリングは説明変数の定義に基づいて行うことは可能であり、今回の調査においてもサンプリングにおける大きな障害はなかった。しかしながら、コントロール群が暴露に関する何の情報もないような場合については、サンプリングのためのロジックを定義することは難しいかもしれない。症例対照研究の場合、サンプリングの対象がアウトカム変数となるため、特にアウトカム発生というフラグを持たない対照群をどのようにサンプルするかについては大きな問題を含んでいるといえる。

<今後の病院情報利用の在り方について>

以上、今回の試験的運用の成果から得られた病院情報システムの二次利用に関する可能性と限界について考察した。以上をかんがみした場合、次のような整備を病院情報システムの基盤に整備していくことが重要であると我々は考える。

- ・ データベースの構造について

データベースについては、すべてのデータを集約化するデータ・ウェアハウスが存在することが望ましいが、必ずしも必須条件ではない。重要なことは、異なるデータベースにあるデータを連結させ、データを一元的に集約させたかのように抽出操作を可能にすることである。岐阜大学が採用している GLASTA のようなシステムを導入することで、これらの問題はクリアされるであろう。問題は、そのようなシステムとデータの相性が悪い場合の対応であるが、データ・ウェアハウスを単体で多くの施設が持つことよりは障害がクリアできる可能性は高い。

- ・ 入力信号について

おそらく、有効な二次利用を行う上で最も重要かつ、バリアの高い部分はデータを入力する時点で、のちに解析を可能とするような信号を発生させることである。これらをあまりに理想的に行おうとすると、現場に大きな負担がかかる。現実的に運用可能であり、さらにそうすることが直接診療機能を向上させることにつながるような入力のフォーマットが必要である。HL7 (Health Level 7) などの国際規格に準拠した信号記録および格納のフォーマットを、現場の負担を最小限にしながらどこまで組み込んでいくことができるかがカギとなるであろう。

- 実行入力について：処置や患者の安静度などについて、どの時点で実行したのかについての実行入力は、デジタルスタンプの形で行わ

れることが望ましい。少なくとも、クリティカル・パスに記載される内容についての実行については、すべからくデジタルスタンプ化することで、二次利用への活用の扉が開けるだけでなく、実際の業務の円滑化にもつながるであろう。医療現場において、実行を行ったときに実行のタイムスタンプが押されるためには、実行確認のためのバーコードリーディングなどが有効である。そのためには、実行確認が患者安全につながる（たとえば薬物の他人への投与の防止など）ような仕組みがあれば、多少の負担があったとしても現場での応用は可能かもしれない。また、看護師の病室への来訪などについては、RF-ID などの運用も将来的には必要になるかもしれない。

- 特に外来診療における基礎情報の標準化：入院データに比較して外来データの二次利用が困難であることは今回の調査でより明確となった。これはDPCデータや、そこに付随する様式1データが存在しないというばかりではなく、他にもさまざまな要素が見られた。それらの要素を修正するためには、第一に外来でのバイタルサイン入力のフォーマット化が要件として必要である。第二には、外来処置に関するタイムスタンプを含めた記号化の整備である。
- 病名について：二次利用を行う場合に大きな障害となる病名の問題

を整理する上では、入院診療記録については資源病名のほかに主病名、およびケースミックスを調整することが可能なリスク調整病名の登録をクリック式で入院時に行うということは現実的には運用できるかもしれない。さらには、のちに提示する退院時サマリを作成する際に整理をすることも有効な方法である。一方、外来における病名については、別途オンゴーイングな問題点として病名登録を行うようなことができれば、いわゆる保険病名と診療病名との区別を行うことができるであろう。

- ▶ 標準化された退院時サマリの運用について：今回、退院時サマリを二次利用することで、その他の診療情報にはない情報を得ることを試みたが、満足のいく結果を得ることはできなかった。これは、現状の電子カルテにおける退院時サマリが二次利用を前提とした構造になっていないことが主たる要因である。退院時サマリに記載されている情報は、テキスト情報以外は他のデータベースに格納されているものである。一方、退院時サマリの一部についてはデータベース構造に基づいた入力を行うことで、診療機能上も質の高いものとなり、さらには職員の負担も軽減させることができる。文書作成システムである「ヤギーシステム[Ⓣ]」は、マイクロソフト アクセスのようにデータベース構造を構築す

ることが可能である。電子カルテシステムにこれらのような文書作成システムを導入している病院は多く、退院時サマリについてもこのような部分を応用することで、二次利用に際して重要な情報を標準的に入力するような環境を構築できる可能性は高いであろう。

- ▶ アウトカムの標準化と、長期アウトカムの把握について：最後に、臨床評価や疫学研究において重要な臨床的アウトカムについては、戦略的な方略が必要となる。術後の生存の追跡や、治療に伴う有害事象の発生などについては、現行の状右方システムで把握することは困難だが、一方で、臨床試験などのノウハウを活用することは可能である。最低限の健康アウトカム情報について、病院単位で標準的なデータ収集と記録を行うシステムを持つことで、議事的に前向きコホート研究データを作成していくことも可能である。これらは、次のステップとして運用上計画することは現実性が高いであろう。

F 結論

今回の二年間の研究事業において、我々は病院情報システムに格納される診療情報等データを二次利用するにあたっての様々な問題点及び示唆を得ることができた。本報告書の内容をもとに、まずは国立病院機構の施設における情報インフラの改善、そして、より広い規模での病院情報の整備に対して情報発信を行うべきと考える。

G 研究発表

1. 論文発表

尾藤誠司、指導医のために プロフェッショナルリズム プロとしての臨床研究 臨床研究指針と実際、日本内科学会雑誌、99 巻 8 号、P1975-1981、2010

小林美亜、古場裕司、尾藤誠司、岡田千春、堀口裕正、三田晃史、伏見清秀、国立病院機構における医療の質評価の取り組み 「医療の質の評価・公表等推進事業」における臨床評価指標に焦点をあてて、病院 69 巻 11 号、885-889、2010

2. 学会発表

M. Kobayashi, S. Bito, C. Okada, K. Fushimi. The development of clinical Indicators for acute care hospitals in Japan. The International Society for Quality in Health Care 28th International Conference. 香港 2011

丹野清美、尾藤誠司、高木安雄、診療費のばらつきと医師の意識の関係 Diagnosis Procedure Combination(DPC) データに基づく分析、第 6 回医療の質・安全学会、東京、2011

小林美亜、尾藤誠司、伏見清秀、医療の質向上につなげるためのクリニカルパスの活用、第 12 回日本クリニカルパス学会学術大会、東京、2011

尾藤誠司、多施設間比較から導き出すベストプラクティスへの変化—国立病院機構の取り組みより—、第 12 回日本クリニカルパス学会学術大会、東京、2011

尾藤誠司、東輝一郎、小河淳、成宮学、山田努 糖尿病外来テンプレートの導入が、糖尿病診療のプロセスと患者アウトカムに与える影響についての前向き研究、第 64 回国立病院総合医学会、福岡、2010

小林美亜、古場裕司、堀口裕正、尾藤誠司、伏見清秀、DPC データを活用した臨床指標算出に係る検討、第 64 回国立病院総合医学会、福岡、2010

古場裕司、小林美亜、堀口裕正、尾藤誠司、伏見清秀、DPC データを活用した病院機能に関する評価方法の検討、第 64 回国立病院総合医学会、福岡、2010

小林美亜、尾藤誠司、伏見清秀、医療の質向上につなげるためのクリニカルパスの活用、第 30 回 医療情報学連合大会 (第 11 回日本医療情報学会学術大会)、静岡、2010

H. 知的所有権の取得状況

なし

二次利用可能性の検討パターンの整理

■用語の定義

☆一次データ：病院の運営管理、診療といった特定の目的のために収集されたデータのこと。

☆二次データ：特定の目的のために収集されたデータ（一次データ）を他の目的で利用する場合のデータのこと。

☆一次利用：特定の目的のために収集されたデータをその目的のために利用すること。

☆二次利用：特定の目的のために収集されたデータを他の目的のために利用すること。

■データ抽出のカテゴリパターン（図 1 参照）

（1）A パターン

・A-1 パターン：「DPC データ・レセプトデータ」（一次利用）の『匿名化が図られたデータ』および「サーベイランス」（一次利用）で『患者 ID の匿名化が図られたデータ』を二次データとして、二次利用をする。

・A-2 パターン：「DPC データ・レセプトデータ」（一次利用）の『匿名化が図られたデータ』および「医療事故報告書」（一次利用）で『患者 ID の匿名化が図られたデータ』を二次データとして、二次利用をする。

（2）B パターン

・「DPC データ・レセプトデータ」（一次利用）の『匿名化が図られたデータ』および「検査データ」（一次利用）で『患者 ID の匿名化が図られたデータ』を二次データとして、二次利用をする。

*検査データ：臨床検査データ、病理報告書、細菌培養検査

（3）C パターン

・C-1 パターン：「DPC データ・レセプトデータ」（一次利用）の『匿名化が図られたデータ』、「診療記録」（一次利用）に記載されているデータのうち、『分析可能なテーブル形式でデータを抽出できるものについて患者 ID の匿名化を図ったデータ』、それを二次データとして、二次利用する。

*診療記録：退院時サマリー、診療・看護記録上のテンプレート、フローシート（温度版）等

・C-2 パターン：「DPC データ・レセプトデータ」（一次利用）の『匿名化が図られたデータ』、「診療記録」（一次利用）から臨床指標算出に係る情報を抽出できるように構造化を図り、抽出した『患者 ID の匿名化が図られたデータ』を二次データとして、二次利用する。

(4) D パターン

・C-1 パターン：「DPC データ・レセプトデータ」(一次利用)の『匿名化が図られたデータ』、「診療記録」(一次利用)の『レビューにより得られたデータ (匿名化処理)』を二次データとして、二次利用する。

これは検証対象としない

一次データ (病院業務)

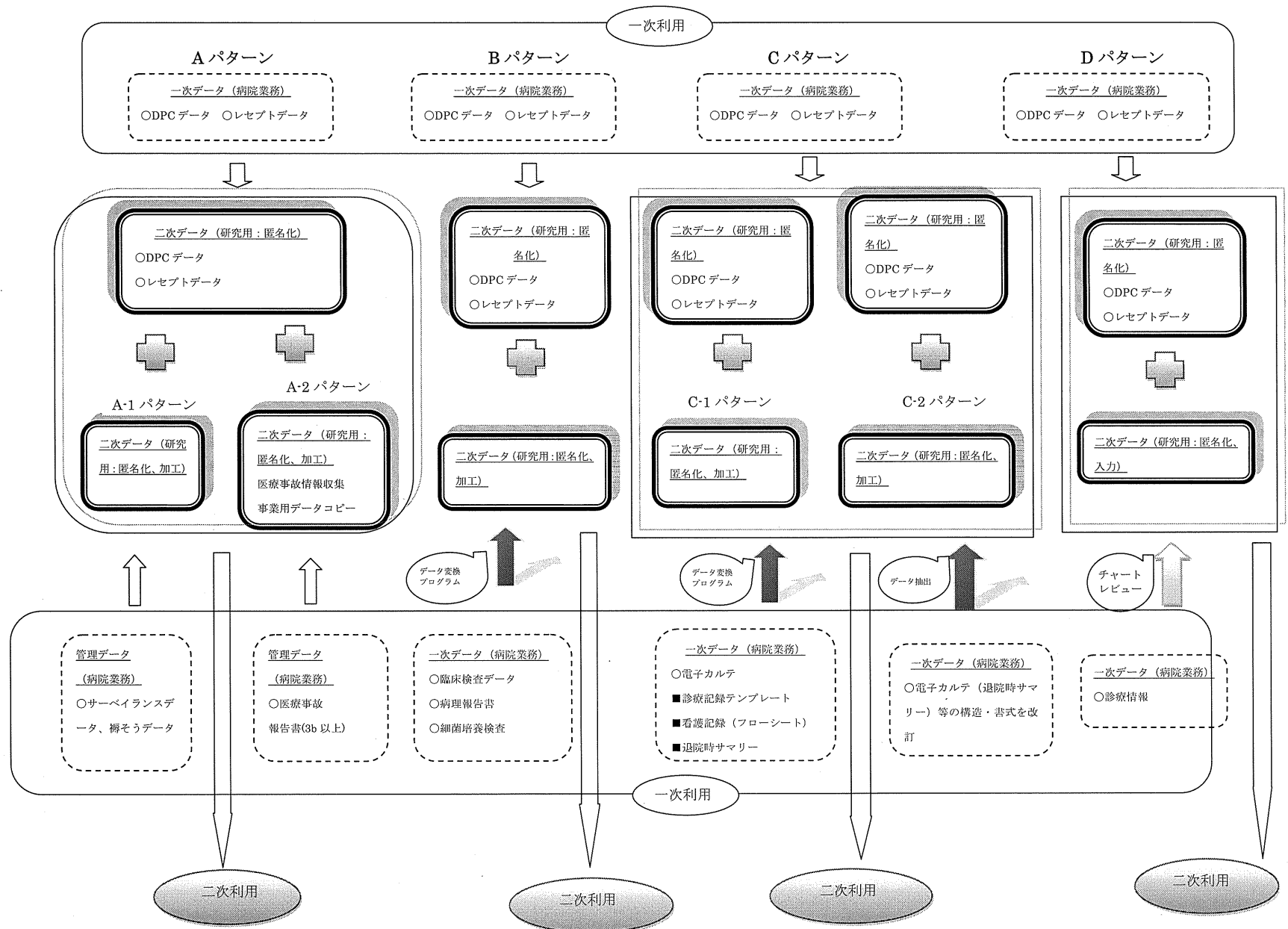
- DPCデータ
- レセプトデータ
- その他

匿名化だけで
ダイレクト使用可能

二次データ (研究用:匿名化)

- DPCデータ
- レセプトデータ

二次利用



「保存された診療データの二次利用適用レベルに準じた、
医療提供プロセスおよびアウトカムの病院横断比較、年
次縦断比較に関する多施設共同研究」

臨床指標の作成用計測マニュアル

目次

1. 高カリウム血症患者における心電図検査の施行率	3
2. 糖尿病患者における血糖値の改善率	5
3. ヘリコバクターピロリの除菌療法の効果	6
4. 気管支喘息のコントロール不良率	8
5. NSAIDs 投与による有害事象の発生率	9
6. 術後呼吸回数の観察実施率	11
7. 手術当日、術前もしくは術中の抗菌薬の投与率(東京医療なし)	12
8. 術後 4 日以内の膀胱留置カテーテルの抜去率	13
9. 術前剃毛の実施率	14
10. 術前プレメディ投薬(東京医療センターなし)	15
11. 術後安静度解除	17
12. SPO ₂ 低下時における胸部レントゲンによる評価の実施率	18
13. 心房細動患者の CHADS2 スコアの測定	20

1. 高カリウム血症患者における心電図検査の施行率

1) 対象データ	入院 様式 1, EF ファイル 検体検査結果データ (カリウム)	
2) 計測期間等	【分母・分子】平成 22 年 10 月 1 日以降入院かつ平成 22 年 10 月 1 日～平成 23 年 9 月 30 日退院患者	
3) 計測対象	分子	分母のうち、高カリウム血症 (6.0mEq/L) が認められた同日、翌日に心電図検査 (D208 心電図検査) が行われた患者数
	分母	50 歳以上 (入院時年齢) で、かつカリウムの血液検査値で高カリウム血症 (6.0mEq/L) が認められた退院患者数

分母の算出方法

以下の基準を全て満たすものを分母とする。

- ① 様式 1 の生年月日、入院年月日より入院時の年齢を求め 50 歳以上のもの。
- ② 様式 1 の入院年月日～退院年月日の期間中カリウム検査が行われ 6.0mEq/L 以上のもの。

分子の算出方法

- ① 6.0mEq/L 以上となった日を求める。

注：入院期間中に 6.0mEq/L 以上となった日が複数日ある場合は初回の 6.0mEq/L 以上になった日とする。

- ② ①の日、あるいはその翌日に D208\$心電図検査 (EF ファイルより以下のレセプト電算コードがあったもの) が行われたものを分子とする。

レセ電算コード	診療行為名称	点数表コード
160068410	E C G 1 2	D2081
160068750	体表ヒス束心電図	D2082
160068850	V C G	D2082
160150650	携帯型発作時心電図記憶伝達装置使用心電図	D2083
160069050	バリストカルジオグラフ	D2084
160069130	2 方向以上加算 (バリストカルジオグラフ)	D2084
160068510	E C G (6 誘導以上)	D2085

*カリウム値

◆検体検査結果データで以下の条件により抽出

- ・検査名称欄が『K』『カリウム』『カリウム(K)』のもの
- ・結果欄に数値が入っているもの
- ・入外区分コードが『入院』か『2』のもの
- ・日付は『検体採取日』を採用（全件『文書日付』と同一）
- ・カリウム値は10以上を除外