

一般口演 | データベース・データウェアハウス

一般口演14

データベース・データウェアハウス

2019年11月23日(土) 09:00 ～ 11:00 J会場 (国際展示場 展示ホール8・特設会場2)

[3-J-1-03] DWHの利用実績からみる電子カルテデータ二次利用状況の変化

○村田 泰三¹、武田 理宏¹、藤井 歩美¹、向井 頼貴¹、真鍋 史朗¹、松村 泰志¹（1. 大阪大学医学部附属病院 医療情報部）

キーワード：DWH, Datamart, Electronic medical record

当院では、1995年より病院情報システムに登録されたデータを分析用データベース（DWH）へ蓄積している。2010年には、オーダ情報のみならず、ペーパーレス電子化カルテとなり、診療記録がすべてDWHの蓄積対象となった。DWHのデータは、2019年6月現在、約28億レコードが約300テーブルに分かれて存在する大規模なデータベースである。これまで、臨床研究支援などの二次利用の取り組みとして、診療情報の抽出代行サービスを実施している。また、DWHから必要な情報を自動抽出し、新たなデータベースを生成するデータマートの仕組みも導入している。

我々は、これらの二次利用サービスを10年以上継続してきた実績がある。そこで、当院での二次利用サービスの実績を見直し、ニーズの変化を捉えるために、現況評価を行った。

抽出代行サービスは、2008年度116件、2012年度500件、2018年度795件であった。2012年度の依頼内容は「統計」が135件(27%)、「症例検索」が326件(65%)、「データ集積」が39件(8%)であった。2018年度の依頼内容は、「統計」が194件(24%)、「症例検索」が418件(52%)、「データ集積」が183件(23%)で、「データ集積」が大きく増加した。

データマートの作成は、2010年度（初年度）が48件、2012年度が150件、2018年度が359件であった。2018年度のデータテーブルを調査すると、複数のデータテーブルを扱ったものが118件(33%)であった。データテーブルの利用頻度は、上位からテンプレートが94件(18%)、薬剤オーダが78件(15%)、入院関連が58件(11%)、問診票が54件(11%)、経過記録が43件(8%)と続いた。近年では疾患レジストリを意識したデータマート作成依頼が増加していた。

DWH の利用実績からみる電子カルテデータ二次利用状況の変化

村田 泰三^{*1}、武田 理宏^{*1}、藤井 歩美^{*1}、向井 頼貴^{*1}、真鍋 史朗^{*1}、松村 泰志^{*1}

^{*1} 大阪大学医学部附属病院 医療情報部

Changes in the secondary use of electronic medical record data based on the usage of DWH

Murata Taizo^{*1}, Takeda Toshihiro^{*1}, Fujii Ayumi^{*1}, Raiki Mukai^{*1}, Manabe Shiro^{*1}, Matsumura Yasushi^{*1}

^{*1} Department of Medical Information Science, Osaka University Hospital

In our hospital, we have accumulated the data, which have been registered in hospital information system, in the data warehouse (DWH) since 1995. We are providing a representative service to extract medical information to support clinical researches using this database. In addition, we have introduced a data mart system that automatically extracts required information from DWH to create a new database. We have continued these secondary use services for over 10 years. Thus, we reviewed the results of secondary use services at our hospital and evaluated current situations in order to grasp the changes in demands.

We performed 116 representative services of data extraction in FY2008, 500 in FY2012, and 795 in FY2018. In FY2012, the numbers of requests in different types were 135 (27%) for “statistics”, 326 (65%) for “case finding”, and 39 (8%) for “data accumulation”. In FY2018, the numbers were 194 (24%) for “statistics”, 418 (52%) for “case finding”, and 183 (23%) for “data accumulation”, indicating that “data accumulation” increased significantly.

We created 48 data marts in FY2010 (first year), 150 in FY2012, and 359 in FY2018. A survey of the use of data tables in FY2018 indicated that multiple data tables had been used in 118 (33%) cases. The types of data table that ranked top in the frequency of use were for template (94, 18%), prescription order (78, 15%), hospitalization (58, 11%), questionnaire (54, 11%), and progress note (43, 8%). Recently, requests for creating disease registries have increased.

Keywords: DWH, Datamart, Electronic medical record

1. はじめに

電子カルテデータの二次利用は、システム導入時に期待する1つのタスクである。当院では、オーダエントリシステムを導入した 1995 年当初にデータウェアハウス(DWH)を構築した。その後、2010 年リプレースで完全ペーパーレス電子カルテとなり、すべての文書を電子化し、入力支援機能のテンプレートの充実化を図り、診療記録の構造化データ登録に取り組んだ。それらのデータは全て DWH に蓄積されている。2019 年 6 月現在、約 28 億レコードが約 300 テーブルに分かれて存在する。1 日約 60 万レコード、1 年に約 2 億レコードが新たに発生する大規模なデータベースである。

我々は、その DWH を用いて、臨床研究支援を中心とした二次利用に取り組んでいる。DWH はシステムベンダが設計するデータベースシステムであり、そのデータを院内でどう活用するかは、病院職員の裁量に任されている。これまで DWH の導入後の活用事例として、患者の待ち時間調査など特定の事例に関する報告はされているが、DWH が病院のどの部署・診療科が利用しており、どんなデータテーブル種が利用されているかなどの要約を報告した例は少ない。

一般的に、DWH の二次利用の方法としては、医療情報部のようなシステム部で抽出代行サービスを実施するか、検索ツールを電子カルテに配置し、医師などのエンドユーザが直接 DWH を操作するかの方法がとられている。当院では、2006 年より前者の抽出代行サービスの運用を開始した。検索依頼数も年々増加しており、この 10 年で、約 8 倍(2008 年度 139 件、2018 年度 795 件)になった。また、2008 年から後者

の検索ツールを配置する方法も導入した。ただし、その検索ツールは、DWH へ直接検索を実行せず、医療情報部がユーザの使いやすい新たなデータベースを作成し、そのデータベースへ検索を実行するWebビューアを提供する仕組み(データマート)とした。データマートは、日次でデータ更新され、エンドユーザは、いつでも電子カルテ端末から検索を行える。

このような医療情報部を介したデータ二次利用の体制をとった結果、電子カルテデータの二次利用の検索要求が医療情報部に集約されるようになった。サービス開始から約 10 年経過したその実績を分析し、二次利用の現況をまとめたので報告する。

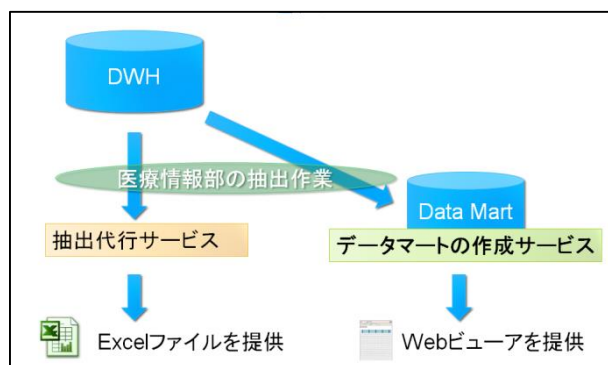


図1 DWH を用いた二次利用サービス

2. 目的

データの二次利用サービスの運用開始し、約 10 年が経過し、抽出代行サービスやデータマートの作成サービスの実績から二次利用のニーズの変化をとらえるために現況評価を行う。

3. 方法

抽出代行サービス、データマートの作成サービスの実績を分析対象とした。抽出代行サービスの分析対象期間は、2008 年度、2012 年度、2018 年度とした。データマート作成サービスの分析対象期間は、2010 年度(初年度)、2012 年度、2018 年度とした。依頼目的別に分類し、件数と割合を算出した。また、部署・診療科別に分類し、件数を比較した。また、データテーブル種別に分類し、その件数と割合を算出し、一度の抽出作業で何種のデータテーブルを利用したかの件数を算出した。

4. 結果

4.1 抽出代行サービスの結果

抽出代行サービスは、2008 年度 116 件、2012 年度 500 件、2018 年度 795 件であった。

4.1.1 目的別の状況

依頼目的を「統計」、「症例検索」、「データ集積」に分類した。「統計」は、外来患者数、入院患者数などの患者数や薬剤投与数などを集計するものとした。「症例検索」は、病名、薬剤、手術、検体検査で異常値であった場合など症例リストを抽出することを目的とするものとした。「データ集積」は、症例リストに対して、検体検査結果や文書のレポート情報など症例リストに対して、アセスメントデータを集積することを目的するものとした。「症例検索」に「データ集積」の目的が加わった複合的な依頼については、「データ集積」へ計上した。

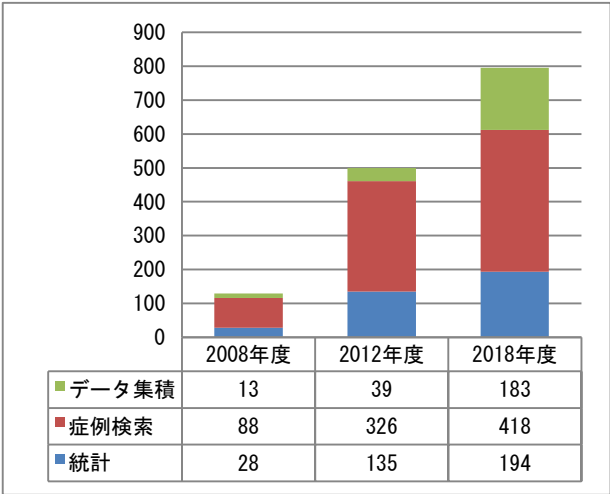


図 2 抽出代行サービスの目的別件数

4.1.2 部署・診療科別の状況

2008 年度、2012 年度、2018 年度に抽出代行サービスを利用した総計は 1424 件であった。その部署・診療科の内訳は、

診療科は 1040 件(73%)、中央診療部は 317 件(22%)、事務部は 67 件(5%)であった。診療科数は、30 診療科、中央診療部数は 14 部署、事務部数は 2 課であった。

年度別の変化をみると、2008 年度は、29 部署・診療科、2012 年度は 41 部署・診療科、2018 年度は 43 部署・診療科から依頼があった。

表 1 抽出代行サービスの部署・診療科別の件数

部署	診療科	2008 年度	2012 年度	2018 年度	総計
診療科	産科婦人科	8	33	65	106
診療科	消化器外科	2	16	65	83
診療科	循環器内科	3	24	52	79
診療科	消化器内科	9	12	53	74
診療科	眼科	5	52	16	73
診療科	血液腫瘍内科	9	33	16	58
診療科	内分泌代謝内科	2	4	49	55
診療科	小児科	3	17	32	52
診療科	呼吸器内科	4	9	31	44
診療科	脳神経外科	3	13	28	44
診療科	腎臓内科	3	17	18	38
診療科	高度救命救急センター	9	23	5	37
診療科	泌尿器科	8	20	6	34
診療科	漢方内科		5	24	29
診療科	神経内科・脳卒中科	3	13	12	28
診療科	神経科精神科	21	3	3	27
診療科	整形外科		15	11	26
診療科	老年高血圧内科	2	3	19	24
診療科	耳鼻咽喉科	1	8	14	23
診療科	皮膚科	4	13	6	23
診療科	麻酔科	1	9	6	16
診療科	心臓血管外科	2	3	9	14
診療科	小児外科	1	5	6	12
診療科	免疫内科	4	5	2	11
診療科	乳腺内分泌外科	2	4	2	8
診療科	集中治療部			7	7
診療科	放射線科		1	6	7
診療科	総合診療部		1	4	5
診療科	形成外科		2		2
診療科	呼吸器外科			1	1
中央診療部	看護部	3	20	49	72
中央診療部	薬剤部	5	28	20	53
中央診療部	医療情報部	2	4	32	38
中央診療部	感染制御部	8	11	16	35
中央診療部	保健医療福祉ネットワーク部		21	11	32
中央診療部	中央QM部	1	13	12	26
中央診療部	栄養マネジメント部		10	12	22
中央診療部	未来医療開発部		5	11	16
中央診療部	オンコロジーセンター		6	2	8
中央診療部	臨床試験部		8		8
中央診療部	移植医療部			2	2
中央診療部	臨床検査部		1	1	2
中央診療部	遺伝子診療部		2		2
中央診療部	病理部			1	1
事務部	医事課	1	8	55	64
事務部	管理課			3	3
総計		129	500	795	1424

4.1.3 データテーブル種の利用状況

抽出代行サービスにおいて、DWH で利用したデータ種について調査した。利用したデータテーブル種数は、2008 年度は、15 種、2012 年度は 26 種、2018 年度は 29 種であった。2008 年から利用頻度の高かった病名登録、入院関連情報、処方オーダ、注射オーダ、外来関連情報、検体検査結果を総計すると 1730 件(66%)であった。2010 年のペーパーレス電子カルテ後のタイミングである 2012 年度から入力テンプレート、経過記録、レジメンオーダ、紹介情報、病理オーダ・レポート、看護関連情報など新たなデータテーブル種が抽出対象データ種に加わった。

また、一度の抽出作業で利用したデータテーブル種数を分類したところ、1種が645件、2種が492件、3種が202件、4種が62件、5種以上が23件であった。

表 2 抽出代行サービスのデータテーブル種別の件数

	2008 年度	2012 年度	2018 年度
病名登録	45	160	230
入院関連情報	38	90	214
処方オーダ	17	106	142
注射オーダ	30	107	125
外来関連情報	22	76	127
検体検査	11	43	147
手術オーダ	12	29	102
医事会計情報	9	42	92
画像生理	12	31	87
文書(サマリ、問診票等)	1	32	52
基本情報(年齢、禁忌など)	1	10	58
入力テンプレート	0	5	35
予約情報	17	7	15
DPC	2	9	26
経過記録	0	9	24
レジメンオーダ	0	0	25
紹介情報	0	3	17
細菌検査	3	8	7
病理オーダ・レポート	0	4	12
看護	0	5	24
食事オーダ	0	7	5
栄養指導	0	4	7
処置	0	2	5
在宅	0	5	0
クリニカルパス	0	2	2
輸血	0	1	3
一般指示	0	0	4
ログイン情報	0	1	2
リハビリ情報	1	0	2
ToDo	0	0	3
総計	221	798	1594

表 3 抽出代行サービスの一度の抽出作業で利用したデータテーブル種数

データ種類数	2008年度	2012年度	2018年度	総計
1種	60	292	293	645
2種	51	136	305	492
3種	14	54	134	202
4種	3	18	41	62
5種	1		14	15
6種			5	5
7種			0	0
8種			2	2
9種			1	1
総計	129	500	795	1424

4.2 データマート作成サービスの結果

データマートの作成は、2010 年度(初年度)が 48 件、2012 年度が 150 件、2018 年度が 358 件であった。データマートは一度作成したデータテーブルを継続して利用しているため、2018 年度 358 件の詳細を報告する。

4.2.1 診療科別の状況

358 件のデータマートを作成した部署・診療科の内訳は、診療科は 240 件(67%)、中央診療部は 109 件(30%)、事務部は 9 件(3%)であった。診療科数は、19 診療科、中央診療部数は 11 部署、事務部数は 1 課、総計 31 部署・診療科であった。上位 5 位は、泌尿器科が 146 件(41%)、中央クオリティマネジメント部が 37 件(10%)、医療情報部 27 件(8%)、整形外科 18 件(5%)、心臓血管外科 16 件(4%)であった。泌尿器科が 146 件と多い理由としては、前立腺がん、膀胱がん患者のリストやカルボプラチン、パクリタキセルなどをはじめとした抗がん剤を投与した患者リストをデータマートとして作成しており、さらに、問診票の OCR 解析や退院時サマリでテンプレートを積極的に利用しているためであった。

表 4 部署・診療科別データマート作成数

部署	診療科	件数	割合
診療科	泌尿器科	146	41%
診療科	整形外科	18	5%
診療科	心臓血管外科	16	4%
診療科	免疫内科	11	3%
診療科	循環器内科	10	3%
診療科	消化器外科	8	2%
診療科	脳神経外科	7	2%
診療科	麻酔科	5	1%
診療科	乳腺内分泌外科	5	1%
診療科	産科婦人科	2	1%
診療科	消化器内科	2	1%
診療科	神経内科	2	1%
診療科	高度救命救急センター	2	1%
診療科	血液腫瘍内科	1	0%
診療科	漢方内科	1	0%
診療科	老年内科	1	0%
診療科	腎臓内科	1	0%
診療科	神経科精神科	1	0%
診療科	呼吸器外科	1	0%
中央診療部	中央クオリティマネジメント部	37	10%
中央診療部	医療情報部	27	8%
中央診療部	看護部	13	4%
中央診療部	感染制御部	9	3%
中央診療部	放射線部	7	2%
中央診療部	薬剤部	5	1%
中央診療部	国際医療センター	4	1%
中央診療部	リハビリテーション部	3	1%
中央診療部	保健医療福祉ネットワーク部	2	1%
中央診療部	栄養部	1	0%
中央診療部	臨床試験部	1	0%
事務部	医事課	9	3%
総計		358	100%

4.2.2 データテーブル種の利用状況

データマートのデータテーブルの利用頻度は、上位からテンプレートが 94 件(18%)、薬剤オーダが 78 件(15%)、入院関連が 58 件(11%)、問診票が 54 件(11%)、経過記録が 43 件(8%)であった。

表 5 データマート作成におけるデータテーブル種

データテーブル種	件数	割合
テンプレート	94	18%
薬剤（処方、注射）	78	15%
入院関連	58	11%
問診票	54	11%
経過記録	43	8%
その他文書	37	7%
病名登録	30	6%
退院時サマリ	21	4%
検体検査	21	4%
手術オーダ	18	4%
外来関連	14	3%
看護	11	2%
画像オーダ、レポート	6	1%
クリニカルパス	6	1%
患者基本（身長、体重等）	5	1%
検査レポート	5	1%
病理オーダ、レポート	5	1%
TODOリスト	3	1%
医事会計	2	0%
リハビリオーダ	2	0%

5 考察

抽出代行サービスの目的別の傾向から、「症例検索」の需要がサービスを開始した当初から高いことがわかる。2018 年度においても 50%以上は、「症例検索」を目的としている。「統計」を目的とする依頼は、25%程度を推移しており、他の検索依頼が増えた状況でも同一割合を保っていることから、アニュアルレポートなどの年次統計だけでなく、その年に新たな作成する統計資料の作成依頼が発生していることがわかる。「データ集積」が年々増加している理由としては、抽出代行サービスが多少複雑な加工作業を請け負っているためだと推察される。例えば、検体検査結果を臨床研究のためにデータ収集する際に、手術の術前、術後 1 週間、退院時、術後 3 ヶ月、術後 1 年、術後 3 年のような定めたイベント日の測定値を集める必要がある。従来の方法では、1 症例ごとにカルテを開き、該当する日付の記録を確認し、転記してデータ集積をすることになる。抽出代行サービスでは、前述した例であってもデータ取得のルールを決めて、機械的にデータを取得するデータ加工を行っており、研究者の転記作業の負担軽減に貢献できていると考えられる。こうした作業実績がロコミで広まっているためか、同一診療科からは似た依頼が寄せられる傾向がある。

また、部署・診療科別の傾向をみると、抽出代行サービスは、中央診療部、事務部よりも診療科からの依頼が多い。上位の診療科を見てみると、産婦人科、消化器外科、循環器内科、消化器内科などの投薬治療を実施する診療科からの依頼が多い傾向がある。「症例検索」の依頼の場合に、病名や薬剤オーダのデータテーブルを条件とするものが多く、投薬治療は手術症例に比べて、より多くの患者に対して治療を実施できる一方で、その投薬治療症例のリスト化を独自に作成するのが困難であるからだと推察する。外科系診療科では、ファイルメーカーによる手術台帳および手術記録を作成している診療科もあり、そういった診療科からの抽出代行サービスの依頼が少ないことが示唆された。

さらに、データテーブル種の利用状況の傾向をみると、病名登録、入院関連情報、処方オーダ、注射オーダ、外来関連情報、検体検査結果、手術オーダ、医事会計情報など 2008 年度から依頼があったものは、依頼件数の増加と比例して、頻度が上がっている。注目すべきは、2012 年度以降に新たに利用するようになったデータテーブル種であり、入力テンプレート、経過記録、レジメンオーダ、紹介情報、病理オーダ・レポート、看護関連情報、指示などは、電子カルテがリプレイスされた結果、利用されることになったデータ種である。看護部では、看護必要度、看護問題、看護計画など看護師がこれまで紙カルテに入力していた情報が電子化されたため、業務評価などに用いるためにデータ解析の需要が生まれている。

また、症例検索の際に、心臓超音波レポートの EF>50 以上などのレポート所見の測定値を条件に設定できるようになったのも電子カルテ化による変化の 1 つである。

利用するデータ種別が増えることにより、一度の抽出作業で利用するデータテーブル種類も増えて、依頼が複雑化している。年間 800 件近くの抽出代行サービスは、2 名のスタッフで実施しており、依頼日から 1 週間までに提供することをサービスの案内に記載しているが、概ね期日が守られている。

次にデータマートの作成の部署・診療科別の傾向をみると、泌尿器科の件数が突出しているが、一方で、抽出代行サービスの利用頻度が下がっていることがわかる。これは、泌尿器科のデータマートは、多種のデータベースを作成したため、エンドユーザがデータマートシステムを利用することで、臨床研究に必要なデータが集積できていることが示唆される。

データマート作成におけるデータテーブル種の傾向を見ると、テンプレートのデータテーブルを利用したものが最も多い。経過記録や文書システムで設定したテンプレートは、あらかじめ、臨床研究のために前向きなデータ集積のために設計したものが多く、それをそのままデータマートへ展開している。最近では、NCD(National Clinical Database)で定義されているデータ集積項目に院内の独自項目を加えるアレンジをして、疾患レジストリのデータベースを作成することがあり、これを実現した診療科の満足度は高い。このように電子カルテデータから診療科独自の疾患データベースを作成することが臨床研究支援の需要の 1 つであることが考えられる。

以上の分析から、当院では抽出代行サービス、データマートの作成サービスにより電子カルテデータの二次利用が有効的に実施できていることがわかる。その一方で、そのサービス利用者は、診療科の偏りもあり、これらのサービスを院内により周知する広報活動が必要であることにも気づかされた。

6 まとめ

抽出代行サービス、データマートの作成サービスの利用実績を利用目的別、データテーブル種別、診療科別に分類し、傾向を分析した。年々、二次利用サービスの実績も増えており、DWH が効果的に活用されている。

今後は、サービス開始当初から多い症例検索だけでなく、疾患レジストリを意識したデータマートの作成を普及させるため、診療科への広報活動をより積極的に行い、臨床研究の中核病院として、一層の研究活動の支援に努めたい。

参考文献

- 1) 村田 泰三, 武田 理宏, 上田 郁奈代ら, ユーザの検索要求からみえるデータウェアハウスの活用の現状, 第 33 回医療情報学

連合大会論文集 2013 :404-407.