

## 分担研究報告書

我が国における臨床データベースの活用性と信頼性の検討

分担研究者 石田 博 (山口大学大学院医学系研究科)

### 要旨

レセプト情報、がん届出情報、電子カルテ情報、および、規定書式によって登録された症例情報等などの臨床データベースについて、その情報種別とともに、その登録割合、ならびに、項目毎の登録情報の登録精度について課題となる症例登録による臨床データベースについて文献等の情報を基に検討した。それぞれのデータベースで登録データの質を検討した論文は少ない状況であったが、NCD 等の症例登録を基にしたものでは、その登録割合や多くの項目で登録精度が高いものであったが、体重や麻酔方法など一部の項目で、また、TNM 分類などでは不一致度の施設による差異が大きいことの報告が見られた。そこで、自施設の登録での差異を確認したところ、同様の項目で誤入力が少ないことが確認され、また、検査結果では、誤入力とともに基準値内外の判断の誤りなど登録方式のよる不一致も見られた。また、症例の病態や行為の特定性についても、レセプト等の傷病名など単独情報の推定では感度が低いケースが多く、薬剤や検査結果などの組み合わせにより感度や陽性適中率を上げることが必要であることを示す報告例が多かった。このような臨床データベースに関する評価情報を事前に確認した上で臨床研究に用いることが重要と考えられた。

### はじめに

費用対効果を含む医療技術評価 (HTA) においては、その治療効果や安全性、費用等の情報として不確実性を最小化するような十分な質のエビデンスが求められるが、これらのエビデンスは限られていることが多い。HTA における追加的有用性のエビデンスの標準法である無作為比較試験 (RCT) を補完するものとして、昨今、リアルワールド・データ (RWD) に基づくエビデンス (RWE) の活用について言及されることが多い。この RWE は、実臨床における無作為化臨床試験である Pragmatic Clinical Trial (PCT) によるものと、日常臨床からの電子カルテ情報、診療報酬情報、種々の疾患や治療毎の症例登録 (レジストリー)、さらには、血糖測定等のユビキタス機器や健康アプリケーションを通して収集されたデータなどを収集・蓄積した RWD から得られるエビデンスに大きくは分けられる。

RWE は、その基となる RWD がどのように収集され、どのように管理、活用されたかによってその質が影響を受けることから、本研究では、PCT を除き、我が国で用いられている代表的な臨床データベースが収集するデータの種別、および、その収集過程からデータの質が

問題となりうる症例登録データベースにおける登録割合と情報の正確性の検討を行った。

## 方法

本研究では、我が国における代表的な臨床データベースの蓄積データの調査、および、蓄積データの臨床的な信頼性検討について文献、Web 調査、および、臨床現場での症例登録状況調査として自施設データによる登録精度の検討を以下の 1~3 のように行った。なお、RWD の質の検討については、登録されたデータが事実であるかの信頼性の検証 (verification) とデータや予測式などが正しく適合しているかの妥当性の確認 (validation) の 2 つの視点のうち、今回は特に前者の verification の観点での検討を主に行った。

1. 代表的な臨床データベースの収集情報の種別とそれらの活用性について、関連論文、および、Web サイト等の情報を基にそれぞれの臨床 DB に蓄積されたデータ種別を調査した。
2. 臨床データベースにおいて最も課題となる登録情報の信頼性の評価について、文献検索ならびに自施設での症例登録における誤登録の状況調査を行った。
  - ① 文献検索：登録精度が最も重要な課題となる直接入力による症例登録データベースにおいて、症例登録の割合（登録漏れがないこと＝完璧性）とその登録の質（元となる情報との一致度）を検討した論文を調査した。
  - ② 自施設での登録精度の検証：上記①の不一致の原因を検討するために自施設での症例登録での不一致症例の検討を行った。具体的には、National Clinical Database (NCD) 消化器外科専門領域における医療水準評価術式様式の登録が求められる胃癌に対する胃全摘術、および、直腸癌に対する低位前方切除術症例で実際に NCD に登録された情報（2014~2015 年：76 症例）を NCD からダウンロードし、カルテ情報を参照基準として、登録データの一致度による検証を行った。尚、NCD の一部の登録書式の検査結果の登録では、NCD が定める基準値内であれば「基準値内」を選択し、範囲外であればその検査結果を入力登録する様式と採っている。
3. レセプトを主体とする臨床データベースにおける患者の正しい病態情報を把握するために検討された論文調査として、DPC データ、MID-NET、NDB を対象とした病態の特定性の評価について検討した。

## 結果

1. 表 1 に我が国における代表的な臨床データベース (DB) をがん登録、診療・介護レセプト関連情報を主体とする公的・民間 DB、次世代医療基盤法の認定事業者の提供する DB、および、学会等の症例登録データベースの 4 つのグループに分け、まとめた。
  - ① がん登録等の推進に関する法律を基とする全国がん登録情報では、全ての病院と都道府県指定の診療所からのがん患者に対して ICD-0-3.1 に基づく患者情報やがんの性状、診断や発見に関する情報、治療に関する情報などが登録され、院内がん登録

では、がん診療連携拠点病院、あるいは、がん診療推進病院等の専門的ながん医療を行う医療機関で全国がん登録との共通項目に標準登録項目、管理項目を加え、より詳細な情報を登録したデータベースである。全国がん登録情報は研究者等にも第三者提供がされているが、院内がん登録情報の全国集計は未だ提供されていない。

- ② 診療レセプトおよび関連情報を主たるデータとする NDB、DPC DB、また、介護レセプト情報・要介護認定情報を主なデータとする介護 DB、および、PMDA による MID-NET が公的なデータベースの代表であり、NDB、DPC DB、介護・要介護情報の 3 者は連結したデータの活用が認められている。特に、NDB、介護 DB などは全国からのレセプトをほぼすべてカバーするもので悉皆性が高いものである。また、NDB は特定健診データも活用することが可能である。DPC DB は DPC 病院での急性期の入院患者についての調査データで、各医療機関から直接、厚労省に 3 ヶ月毎に定期的に提出されたデータを収集したものである。

民間の代表的なデータベースとしては、JMDC database、Medical Data Vision (MDV) データベースなどがある。いずれも健康保険組合や契約する医療機関から取得した診療報酬請求のレセプトデータなど医事関連データを主たる情報とし、一部の医療機関から検査結果などの電子カルテ情報も取得、活用可能としている。

さらに、PMDA(日本医薬品医療機器機構) の MID-NET は医薬品の安全対策等を推進するために、参加医療機関 (10 拠点、23 医療機関) のレセプトデータ、DPC 調査データに加え、検査結果を中心とする電子カルテシステムデータが収集されているものである。特に処方、注射オーダーや検査結果の収集については、各施設でのコードの標準化などの品質管理が重視されている。

- ③ 次世代医療基盤法による臨床データベースは、令和 4 年 3 月時点で、2 つの認定匿名加工医療情報作成事業者がデータの提供を行っている。ライフデータイニシアティブ(LDI)は主として医療機関から電子カルテ情報、DPC 調査情報、レセプト情報を取得して匿名化情報を提供するものであり、日本医師会医療情報管理機構(J-MIMO)は医療機関からの情報に加えて、一部の自治体や保険者からの医療・介護レセプト情報、健診情報などを収集して提供するとしている。
- ④ 主として学会を事業者として症例登録を行うもので、代表的なものとして外科系学会のそれぞれの専門医制度とも連動し共通プラットフォームを基盤とした NCD、日本糖尿病学会の J-DREAMS、日本腎臓病学会の J-CKD-DB などがある。NCD は、基本的に Web での登録書式 (Case Report Form: CRF) への直接登録であり日本外科学会、日本消化器外科学会、日本胸部外科学会、日本小児外科学会等の多くの外科系専門医領域別の手術症例を中心とした登録ばかりでなく、胃癌や肝癌登録などにも活用されている。一方、J-DREAMS は電子カルテに糖尿病診療用 CRF を導入し、日常診療において登録された情報と SS-MIX2 標準ストレージの処方や検査情報をそれぞれの施設の SS-MIX2 ストレージより多目的臨床データ登録システム MCDRS

(<http://mcdrs.jp/>)を用いて収集している。また、J-CKD-DB は、慢性腎臓病患者を対象に、関連の病名、処方、検査情報などと同じく MCDRS を用いて参加医療機関の SS-MIX2 標準ストレージから収集している。

以上の 4 グループの内、②、③の臨床データベースについては、大学等の研究機関ばかりでなく、民間の事業者でも一定の手続後の審査結果によって活用が可能な状況にある。

## 2. 登録データの質の検討

### ① 登録症例の完璧性と登録項目の一致度の評価：

症例登録事業の中で一般的な臨床情報の登録を行う NCD 関連の論文が 5 件と DPC 調査データを対象とした論文が 1 件確認され、表 2 にまとめた。

NCD 関連のうち、2 件は先天性心臓血管手術症例データベースであり、消化器外科専門領域の症例登録を対象としたものが 2 件、および食道手術と肺手術症例の登録を対象としたものが 1 件であった。DPC 調査データを対象としたものは院内がん登録への活用性について検討したものであった。

先天性心臓血管手術症例における症例登録の割合の検討はいずれも施設訪問による手術記録の確認によるもので 99%と高率であった。一方、**regional government report** との比較では、対象が肺切除、食道手術といった術式に限定されるが、登録開始年である 2011 年に全体の 90~95%の症例が登録されており、高い登録率が示さ、登録症例の良好な「代表性」が示された。一方、個々の項目で本来の情報と一致しているかとの観点では、どの検討でもほとんどが 9 割は超える一致率であった。

項目別の登録情報については、参照基準である手術記録等の比較で多く項目で高い一致度が示されたが、体重や麻酔法などの一部の項目では一致率が低めとなる傾向であった。また、Okuyama らの院内がん登録を参照基準とした悪性腫瘍の TNM 分類情報の調査では、DPC 調査データにおいて全体でも T および N については 9 割前後の一致率と低めであり、また、施設別のバラツキが大きい結果であった。

### ② 自施設における検討：

2014~2015 年の NCD 消化器外科専門領域消化器外科医療水準術式対象として登録された胃全摘術、低位前方切除術施行 76 症例におけるデータについての電子カルテ上の情報との不一致率を確認した結果を表 3 に示す。①の検討でも一致率が低い項目として示された麻酔法や TNM 分類に加え、検査結果で不一致が比較的高い割合で確認される結果であった。この原因としては、TNM 分類等ではタイプエラーによるものが多いと考えられ、特に、体重や検査値など数値等の入力間違いが起りやすいことが考えられた。また、検査結果では自施設内と NCD とで基準値の上・

下限が異なることもあり基準値内か異常値かの判断の誤りが多く見られ、また、確認漏れと考えられる「未検査」とするものが多いことが影響した結果と考えられた。

### 3. 臨床データベースによる死亡転帰や病名等の病態、行為数の特定性の評価

表4に、DPCレセプトを含むレセプト情報、あるいは、SS-MIX2標準ストレージからの検査結果を参照することで、死亡や癌や糖尿病、Charlson併存疾患等の患者の特定、および、処置行為数、検査実施の記録の有無についての特定性、確認を感度、特異度、あるいは陽性適中率で評価を行った論文についてまとめた。いずれも特異度は高いが、感度は低い結果であった。そのために、その疾患に使われる薬剤や検査結果などを定義に加えることでどの程度の感度が高くなるかについて評価がなされているものも多く見られた。また、NDBオープンデータにおける抜歯数に対する検討では、全国の歯科医師への期間調査からの推定値との比較で、年齢によっては有意な差異のあるものの、その差異は0.05歯/年間と小さかったことなど、NDBオープンデータの活用性を評価している。

## 考察

我が国でもがん登録情報、レセプト情報をもととしたNDBや外科学会で専門医制度と連携したNCDなど悉皆性の極めて高く大規模な臨床データから<sup>1</sup>、希少疾患や学会等が運営するような小規模なものまで様々な臨床データベースが構築され、運用されている<sup>2</sup>。それらの臨床データベースを活用する際には、その蓄積されたデータの特性として、どのような臨床情報がどのように収集、管理され、どのような質の情報として提供されているか、また、そのデータがいわゆる代表性を有するような集団から収集されたものかといった観点での確認が重要である<sup>3</sup>。

臨床データベースを構成するデータ種別として、診療報酬請求に必要となる傷病名や用いられた医薬品、さらに、手術や主要な処置行為は、レセプト電算コードなどで標準化された情報として、また、収益に直結するものとして医師の確認の上、提出されるものであり、比較的登録における誤りの少ない情報と考えられるデータである。しかし、検査等の実施のために保険病名と称される傷病名が登録されたり、逆に、診療報酬とは直接には結びつかない、あるいは、その診療には直接、関連のない病態が存在しても傷病名として登録されないことも少なくないこと、また、死亡や病態が解決していてもレセプト上は転帰が入らず継続されるなど傷病名等の病態情報は他の情報と比較し精度の低い情報が収集されている状況である。一方、DPC DBは、急性期入院における包括支払いの情報でありその病名は信頼性が高いが、併存症や合併症については診療報酬請求に関連する副傷病名以外の傷病名が同様に登録されないことが少なくないことが想定される。このような背景から多くの論文で、正しい病態の特定についての報告がなされている。特に併存症については治療効果等の解析の際に問題となる交絡にも関連し、重要な情報であることからその特定の精度向上は今後

も継続する課題である。

さらに、電子カルテ情報については、レセプトなどの医事関連の情報に加えて、検査情報が活用できる利点がある。しかし、病院内での検査情報は、それぞれの施設で固有のオーダコードが用いられていることから、臨床データベースとして多施設の情報をを用いる場合には、検査項目や医薬品毎に標準化コードを誤りなく **Mapping** して用いることが重要である。そのような事から **MID-NET** では、別途、**MID-NET** マスターを作成し、それを施設毎マスターの中核として管理するなど標準化コードの統一管理を重視した信頼性の高い運用がなされている。

現時点では、これらの **DPC** 調査情報などの診療報酬請求関連情報、あるいは、**SS-MIX2** 標準ストレージに含まれないような臨床情報を電子カルテから取得することは難しいこともあり、症例登録においては基本情報や入退院情報、検査値以外の情報の直接の入力が課されていることが多く、その症例の取りこぼしや登録された情報の精度が **RWE** の質にも影響する。そのために、カルテ上の病歴等の情報と一致した情報の登録がなされているかの登録精度の検討が重要であるが、今回、調査した中では、そのような検証報告を論文等の形で報告している症例登録事業は少ない状況にあった。その中で、**NCD** 等では現地のよる手術情報等の調査をもとにした検証の結果で、その高い症例登録率、ならびに項目の一致率が報告されている。しかし、その中でも体重や **TNM** 分類など 1 割程度の差異があるものがあり、自施設例からみると、その多くがタイピングや判断の誤り、確認漏れなどであり、さらに、がん登録などでは、施設間差も示されていることなどからそれらの登録精度に関わる状況を理解した上で用いることが重要と考えられた。

一方、これらの情報項目の中でも検査結果などの登録時の入力で間違えることが多いものであるが、電子カルテにある情報を **SS-MIX2** 標準ストレージなどを介して、自動的に取りこむことで結果入力の負荷軽減と誤入力予防になることから、今後の症例登録においては、**J-DREAMS** 等に用いられている **MCDRS** のような **SS-MIX2** 情報を直接、連携するシステムが期待される<sup>4</sup>。

さらに、用いるデータの「代表性」の観点では、**NDB** など、対象となる人口のほぼ **100%** に近い登録がなされているものを筆頭に、レセプトベースのものは大規模なものが多いが、情報取得が健保組合のみで国保などを含まないものは、高齢者や若年者の患者が抜けるなどの問題がある。また、学会の症例登録である **J-DREAMS** や **J-CKD-DB** などでは大学病院や大規模病院の専門外来の患者が主要な登録対象であったりすることから、糖尿病や慢性腎臓病患者の代表性についての考慮が必要と考えられた。

表1 代表的な臨床データベースとその情報内容

臨床データベース	がん登録 (全国・院内) <sup>1</sup>	NDB (レセプト情報・ 特定健診等情報データ ベース) <sup>2</sup>	DPC DB <sup>3</sup>	介護DB <sup>4</sup>	MID-NET <sup>5</sup>	商用データベース (JMDC <sup>6</sup> ・MDV <sup>7</sup> 等)	次世代医療基盤法による特定 認定事業者* (ライフデータ イニシアティブ <sup>8</sup> ・J- MIMO <sup>9</sup> 等)	NCD <sup>10</sup>	J-DREAMS <sup>11</sup> / J- CKD-DB <sup>12</sup>
がん登録情報	がん登録情報：全国がん (患者番号、性別、診断時年齢、住所・局在 (ICD-O-3)、診断名、ICD-10、診断根拠、診断年月日、発見経緯、進展度、治療 (外科・鏡視下治療・内視鏡的・観血的・放射線治療・化学療法・薬物治療・原死因・死亡年月等	○ 院内がん登録では臨床病期、TNM分類等が加わる							
保険情報等	医療レセプト情報：傷病名/診療行為 (処方・手術・処置・検査・指導等)・医薬品・特定器材等 DPC調査情報：様式1 (患者属性や病態等の情報)、Hファイル (重症度、医療・看護必要度情報)、E-Fファイル (出来高情報：診療明細・行為明細で入外を含む)、包括支払い情報等 介護レセプト・要介護認定情報・LIFE情報：給付実績情報 (基本情報/明細情報/緊急時施設療養/特定診療費/食事費/居宅サービス等)・要介護認定情報・LIFE情報 (利用者情報・科学的介護推進情報等	○	○ (集計表)		○	○	○	△	
健診情報	特定健診・特定保健指導データ：受診情報、受診者情報、健診・問診情報・保健指導レベル、特定保健指導のポイント数等	○		○		○	J-MIMO		
電子カルテ標準化情報	SS-MIX2標準化ストレージ情報：患者基本情報 (病名・アレルギー情報を含む)、入退院、処方・注射、検体検査・結果、放射線検査、内視鏡検査、生理検査のオーダ・実施情報				○	△	○		○
電子カルテ非構造化情報を含む	電子カルテ情報：患者基本・患者プロフィール (アレルギー・嗜好歴等)、処方・注射、検査結果、内視鏡・放射線レポート・手術サマリ等						○		
書式登録情報	書式ベースの登録情報：患者基本情報、疾患、治療情報、関連情報など							専門領域別手術・処置患者 (消化器外科・循環器内科・外科・呼吸器外科等)、胃癌・乳癌・肝癌等患者 (非手術例を含む)	J-DREAMS：糖尿病

\*認定匿名加工医療情報作成事業者

\*\*一般財団法人 日本医師会医療情報管理機構 (J-MIMO、Japan Medical Association Medical-Information Management Organization)

1. がん登録 ([https://ganjoho.jp/med\\_pro/cancer\\_control/can\\_reg/index.html](https://ganjoho.jp/med_pro/cancer_control/can_reg/index.html))
2. 匿名レセプト情報・匿名特定健診等情報の提供に関するホームページ ([https://www.mhlw.go.jp/stf/seisakunitsuite/bunya/kenkou\\_iryuu/iryuhoken/reseputo/index.html](https://www.mhlw.go.jp/stf/seisakunitsuite/bunya/kenkou_iryuu/iryuhoken/reseputo/index.html))
3. 匿名診療等関連情報の提供に関するホームページ ([https://www.mhlw.go.jp/stf/seisakunitsuite/bunya/kenkou\\_iryuu/iryuhoken/dpc/index.html](https://www.mhlw.go.jp/stf/seisakunitsuite/bunya/kenkou_iryuu/iryuhoken/dpc/index.html))
4. 匿名介護情報等の提供について ([https://www.mhlw.go.jp/stf/shingi2/0000198094\\_00033.html](https://www.mhlw.go.jp/stf/shingi2/0000198094_00033.html))

5. MID-NET(<https://www.pmda.go.jp/safety/mid-net/0001.html>)
6. JMDC Claims Database(<https://www.jmdc.co.jp/jmdc-claims-database/>)
7. MDV(<https://www.mdv.co.jp/ebm/>)
8. ライフデータニシアティブ (<https://www.ldi.or.jp/>)
9. J-MIMO(<https://www.j-mimo.or.jp/>)
10. NCD(<http://www.ncd.or.jp/>)
11. J-DREAMS (<http://jdreams.jp/>)
12. J-CKD-DB( <http://j-ckd-db.jp/>)

表2 代表的臨床データベースにおける登録データの登録割合・一致度

対象グループ	症例登録					DPC調査データ
PMID	24403354	29544411	30088123	30406495	34095719	31947617
著者(第1)	Tomotaki A	Takahashi A	Tomotaki A	Kanaji S	Hasegawa H	Okuyama A
出版年	2013	2018	2019	2019	2020	2020
質の指標/活用性指標	登録割合の完璧性・項目内容の一致度	登録割合の完璧性・項目内容の一致度	登録割合の完璧性・項目内容の一致度	登録内容の一致度	登録内容の一致度	登録データの精度
対象DB	JACCVSDデータ (2008年8月~2009年12月)	JACCVSDデータ (2010年8月~2012年12月)	NCD登録データ (2011年登録症例)	NCD：消化器外科専門領域 (2014年~2015年)	NCD：消化器外科専門領域 (2016年~2018年)	DPC調査データ (2012年10月~2014年12月末)
特定または検証内容	手術登録割合の完璧性・登録内容(25項目)の正確性	手術登録割合の完璧性・登録内容(25項目)の正確性	手術登録割合の完璧性・登録内容(8項目)の正確性	登録内容の正確性	登録内容の正確性(現地と遠隔での検証)	登録項目の完全性・正確性
参照基準	6施設の訪問による手術記録チェック	9施設の訪問による手術記録チェック	厚生局届出の手術症例数(肺・食道の手術) 19施設の訪問による手術記録チェック	17施設(大学病院・高症例施設)の訪問による手術・麻酔記録・退院サマリのチェック	遠隔対応：88施設 現地訪問：48施設	院内がん
精度指標	登録割合・一致度	登録割合・一致度	登録割合・一致度	一致度	一致度	不明率・一致度
対象数	1279手術・項目の一致度(129名)	1928件の手術の登録数・項目の一致度(191名)	地域白書(regional government report)との比較：肺切除術(48716症例),食道手術(8399例) 施設訪問での比較：手術記録(2829名)項目の一致度(616)	338症例(20症例×17施設-2例)	遠隔対応：80,776人 現地訪問：44,160人	231医療機関：180,399人 胃癌：48,297人 肺癌：43,310人 結腸/直腸癌：38,881人 乳癌：36,097人 肝癌：13,814人
結果	1. 手術登録数の登録割合：1266/1279(99%) 2. 登録内容の一致度： ・症例別：中央値91%(4分位86%~95%) ・項目別：43%(気管挿管の日時)、身長(0.87)、体重(0.83)~100%(死亡日・術後30日・90日状態等) →登録割合(登録の完璧性)は高く、項目別の登録内容の一致度は許容範囲であった。	1. 手術登録数の登録割合：1910/1928(99.1%) 2. 登録内容の一致度： ・項目別：75%(体重)~100%(性別・血型・退院時状態、術後30日・90日状態等) ・術後合併症：特異度 $\geq$ 0.99と高いが、感度は創部感染70%、再手術67%と低い項目があった。 →登録割合(完璧性)は高く、また、項目別にはカテゴリー項目は高い精度で登録されていたが、術後合併症に一部、感度が低いものがあり、継続的なデータの質向上が必要。	・NCD登録症例の登録割合 1. regional government reportとの比較→国内で実施された90-95%がNCDに登録されていた。 肺切除術：46143/48716症例(94.7%) 食道手術：7494/8399例(89.2%) 2. 施設訪問での調査 ・NCDに登録された2829症例のうち2783例(98.4%)が正しく登録されていた。 ・登録内容の一致度：全8項目97.8% 項目別：95%(入院日)~99.4%(退院時状態) →カルテ記録調査からNCD登録の登録精度は高い結果であった。	入力的一致度 ・全体：17施設の45項目の施設中央値は98.2%(範囲：96.4~99.9%) ・術前情報・術中情報(22項目)：体重(85.8%)~腫瘍の性状(100%) ・術後情報(21項目)：再入院(98.22%)~末梢神経障害(100%) 術後合併症で多い創部感染では98%台であった。 不一致の主な原因はタイプエラーである。 →消化器外科領域での登録情報の精度は高いことが示された。	・現地調査、および、遠隔での一致率は、それぞれ、98%、98.2%であった。 ほとんどの項目は一致度>95%<95%の項目は、創傷クラス(92.7%)、体重(89.3%)、麻酔法(90.5%)、ASAクラス(87.2%)であった。 →NCDデータは高い精度で信頼しうる。また、データ検証(Verification)は遠隔対応でも違いは少なかった。	1. 情報の不明/欠落率 身長/体重：2.3% 喫煙index：10.5% 入院時ADL：0.3% 退院時ADL：0.6% 喫煙Indexでは施設間差が大きい(0~93%) 2. TNM分類の一致度 T：86.8% N：91.3% M：95.8% TNM全て：80.6% ・施設別的一致度：40.8%~100% →身長/体重、ADLでは明らかな問題はなかったが、TNM分類に差異が認められた。

1. Tomotaki A, et al. Results of data verification of the Japan congenital cardiovascular database, 2008 to 2009. World J Pediatr Congenit Heart Surg. 2014;5(1):47-53. PMID: 24403354.

2. Takahashi A, et al. Verification of Data Accuracy in Japan Congenital Cardiovascular Surgery Database Including Its Postprocedural Complication Reports. *World J Pediatr Congenit Heart Surg.* 2018;9(2):150-156.. PMID: 29544411.
3. Tomotaki A, et al. Evaluating the quality of data from the Japanese National Clinical Database 2011 via a comparison with regional government report data and medical charts. *Surg Today.* 2019;49(1):65-71. PMID: 30088123.
4. Kanaji S et al. Initial verification of data from a clinical database of gastroenterological surgery in Japan. *Surg Today.* 2019;49(4):328-333. PMID: 30406495.
5. Hasegawa H et al. Validation of data quality in a nationwide gastroenterological surgical database: The National Clinical Database site-visit and remote audits, 2016-2018. *Ann Gastroenterol Surg.* 2020;5(3):296-303. PMID: 34095719
6. Okuyama A, Higashi T. Usability of Clinical Information in Discharge Summary Data in the Diagnosis Procedure Combination Survey for Cancer Patients. *Int J Environ Res Public Health.* 2020;17(2):521. PMID: 31947617

表 3-1 当院における NCD 消化器外科専門領域 医療水準術式登録における一致率—NCD データとの比較  
(胃全摘術・低位前方切除術 2014～2015 年登録 76 例)

項目	Kanaji S <sup>1</sup>		Hasegawa H <sup>2</sup>			自施設				
	施設訪問	全体	施設訪問	遠隔	臨床	検査	%	基準値⇔ 異常値*	誤入力**	その他***
生年月	98.8	99.2	99.1	99.2	100	白血球	94.7	2	0	2
性別	99.4	99.9	99.7	99.9	100	ヘモグロビン	90.8	4	1	2
入院日	97.9	97	96.8	97.1	98.7	ヘマトクリット	90.8	4	1	2
術前診断	98.2	99	99.2	99	94.7	血小板	86.8	8	0	2
退院日	98.8	97.6	97.1	97.8	98.7	アルブミン	92.1	4	0	2
喫煙歴 (Pack-Year)	96.7	95.9	95.3	96.3	97.4	総ビリルビン	96.1	1	0	2
透析	99.7	99.9	100	99.9	100	AST	93.4	3	0	2
悪性腫瘍の部位	99.4	99.1	99.4	98.9	100	ALT	92.1	4	0	2
身長	99.1	98.1	97.1	98.6	98.7	ALP	92.1	4	0	2
体重	85.8	89.3	87.9	90	94.7	尿素窒素	93.4	3	0	2
麻酔方法	89.4	90.5	91.1	90.1	84.2	クレアチニン	89.5	0	6	2
T					97.4	血清ナトリウム	94.7	2	0	2
M					90.8	HbA1c	90.8	0	5	2
N					96.1	CRP	76.3	16	0	2
心停止		100	100	100	100.00	PT INR	93.4	3	0	2
術中/術後輸血		98.2	98.1	98.2	97.4	APTT	65.8	23	1	2
予期せぬ再手術		98.9	99.4	98.6	100					

NCD 症例登録における登録項目の精度について、自施設の調査と比較したものである。検査項目については 2 つの論文では示されていないため、自施設での調査情報のみとなっているが、その不一致例での状況を追加している。基準値内か異常値かの判定に誤りがある例が多いが、基準値が示されていない項目では数値の誤入力と思われる例が多く認められた。

<sup>1</sup>Kanaji S, et al. Initial verification of data from a clinical database of gastroenterological surgery in Japan. Surg Today. 2019;49(4):328-333. 7. PMID: 30406495.

<sup>2</sup>Hasegawa H, et al.. Validation of data quality in a nationwide gastroenterological surgical database: The National Clinical Database site-visit and remote audits, 2016-2018. Ann Gastroenterol Surg. 2020;5(3):296-303. PMID: 34095719

\*基準値、異常値の判断を間違えているもの

\*\*入力間違い、「未検査」と登録したもの

\*\*\*異なる日付の結果を記載(2名)

表 3-2 TNM 分類不一致例（7 例）の詳細

病期 分類	入力値の不一致		
	症例	退院時サマリ等	登録情報
T分類	①	T4b	T3
	②	T1a	T3a
N分類	①	N1	N2a
	②	N0	N2
	③	N1	N3a
	④	N3	N2a
	⑤	N2	N1a
	⑥	N1	N0
	⑦	N0	N3
M分類	①	M1	M0
	②	M1	M0
	③	M0	M1a

TNM 分類で差異のあった 12 例についてその内容を示す。

表 4 臨床データベースによる死亡転帰や病名等の病態、行為数の特定性の評価

PMID	23741526	25744458	28142051	34776478	34495957
著者(第1)	Ooba N	Sato I	Yamana H	Suzuki S	Watanabe H
出版年	2013	2015	2016	2021	2021
質の指標/活用性指標	レセプトにおける死亡転帰の特定性	乳癌患者の特定性	併存症・処置の特定性、検査実施記録の一致性	歯科処置の特定性	疾患の特定性
対象DB	JMDC:レセプト情報 (2005年1月~2009年12月)	施設DPCレセプト (2011年1月~12月)	DPCデータ+SS-MIX (2014年4月~2015年3月)	NDB Open data (2017年4月~2018年3月)	MID-NETデータ レセプトデータ+SS-MIX2データ (2007年4月~2011年12月)
特定または検証内容	各種定義による死亡特定	各種定義による乳癌患者の特定	16のCharlson併存症の有無・10種類の処置の有無・13の検査における一緻度	抜歯数の特定	疾患の特定 (糖尿病・甲状腺機能亢進症・高脂血症・急性腎不全)
参照基準	保険の加入状況 (enrollment)	院内がん登録	カルテ確認	歯科医師へのアンケート調査 (2018年6月4~10日) (回答率: 2345/5250=44.8)	検査結果値による判定基準
精度指標	感度・特異度・陽性適中率	感度・特異度・陽性適中率	感度・特異度・陽性/陰性適中率・一緻度	推定抜歯数比較	陽性適中率
対象数	20~74歳の患者195,193人 死亡転帰者: 680名	聖路加国際病院入院患者 (2011年1月~12月) 乳癌633名/非乳癌50,056名	国立病院機構の4つの急性期病院に2014年度に入退院した (のべ) 315名	日本全体における抜歯数	6医療機関: SS-MIX2 2医療機関: 電子カルテ記録+レセプトデータ
結果	1. レセプト上の死亡転帰: 感度 58%、特異度99.99%、陽性適中率 95.6% 2. レセプトの中で長期にわたるZombieレセプトを除く: 感度 58%、特異度99.99%、陽性適中率 95.6% 3. 重症定義をCharlson併存症のCCI(commorbidity index)I $\geq$ 6で定義した場合、感度31.6%、特異度 99.96%、陽性適中率74.1%  →レセプトによる死亡転帰の特定における感度は約60%だが、特異度は非常に高く、陽性適中率も高い	乳癌の特定をDPCレセプトに含まれる9つのコード (傷病名、手術、腫瘍マーカー、化学療法等) の組み合わせによる14個の定義での感度、特異度は、感度: 32.4~98.7、特異度 99.3~99.9であり、最も高いPPVは 87.3%であった。 →傷病名+手術、化学療法、薬剤、放射線治療のいずれかが乳癌に該当するコードの場合に最も高率に乳癌を特定できた	1. HIVを除くCharlson 併存症の疾患別に感度は0(半身麻痺) ~ 83.5%、特異度は>96%と高く、陽性適中率は半身麻痺を除いて50~100%であった。 2. 処置行為の特定は、感度が最も低いのはパルスオキシメーターで 21.1%で特異度も最も低く、88.5%であった。それ以外は、感度60%以上、特異度は90%以上であった。 3. 検査の実施の一緻度は96%以上であった。  →DPCデータによる診断、処置の特定性、SS-MIXによる検査の特定は一般的に高い結果であった。	年齢階級別・性別での、歯科医師アンケートから推定される抜歯数との比較: NDB open データによる抜歯数は、アンケート調査からの推定数の信頼区間よりも55~59歳、70~74歳で少なく、25~29歳、40~44歳で多いという違いはあるが、全体では 0.05歯/年未満の差であった。  →NDBは日本全体での抜歯行為の特定において信頼しうるデータソースであると示唆される。	陽性適中率は以下の通り 糖尿病新規患者: レセプトデータ: ~40.7% SS-MIX2: ~44.7% 高脂血症新規患者: レセプトデータ: ~73.3% SS-MIX2: ~63.8% 甲状腺機能亢進症新規患者: レセプトデータ: ~72.4% SS-MIX2: ~70.0% 急性腎不全: SS-MIX2: 過去で高クレアチニン患者を除くと89.5%に上昇 →疾患の特定にその疾患特有の状態を含めることが重要であり、臨床検査結果も含めると非常に有用である。

## 参考文献

1. Hiramatsu K, Barrett A, Miyata Y; PhRMA Japan Medical Affairs Committee Working Group 1. Current Status, Challenges, and Future Perspectives of Real-World Data and Real-World Evidence in Japan. *Drugs Real World Outcomes*. 2021 Dec;8(4):459-480.PMID:34148219
2. 日本薬剤疫学会 データベース一覧 : <https://sites.google.com/view/jspe-database-ja2020/> (2022年3月 access)
3. Wang SV, Schneeweiss S, Berger ML, Brown J, de Vries F, et al.; joint ISPE-ISPOR Special Task Force on Real World Evidence in Health Care Decision Making. Reporting to Improve Reproducibility and Facilitate Validity Assessment for Healthcare Database Studies V1.0 *Value Health*. 2017;20(8):1009-1022. PMID: 28964431.
4. 石田 博 電子カルテ情報によるNCD 症例登録支援 *Precision Medicine* 2020 3(9):800-803