

理、情報システムのセキュリティや継続性を保証するシステム管理、という分業の下で事業を進めている。

3. 計画管理

1) データ入力担当者

情報システム自体が完成しても、実際にデータが入力されなければ臨床データベースは稼働しない。一方で臨床データベースは、名称が示すとおり専門的な臨床定義を伴う項目入力を伴うため、入力の手間は少なくない。米国 ACS ではデータベースの参加に当たって、データ入力の担当者の教育プログラムをセットにしたプログラムを実施している。

一方で日本においては、各臨床科別にデータ入力担当者の人件費を財源として確保することは容易ではなく、臨床現場の医師が入力を担当するケースも少なくない。継続的かつ、質の高いデータを共有する上で、今後はデータ入力や検証を支援するための人員と財源を確保することも重要な課題である。

2) ソフトウェア開発

米国の臨床データベース事業では、各部門が年間数十万の費用を支払って、入力用のソフトウェアを購入して、データベース事業に参加するケースが多い。この場合運営側は、入力項目と定義を決めるのみで、ソフトウェア開発自体を行う必要がない。

一方で日本においては、小規模施設が多いため各診療科単位で同等額のソフトウェアを各施設診療科単位で購入することは、財源の観点から容易ではない。この場合、最低限の機能のソフトウェアについては、運営側が開発し、無償で配布することが重要な要素となる。したがって臨床データベースの運営主体は、管理・運営費の中にソフトウェア開発・維持の経費を含めることが必要となる。

3) 入力項目、入力体制の検討

現実的な進行を考える上で参加施設の選定も考慮すべき事項である。臨床データベースは多くの

場合、比較的大きな規模の大学病院が中心となってプロジェクトを立ち上げることが多く、入力項目の基準や入力体制はそれら施設を基準にして検討されることが多い。しかしながら、医療の質向上を目的とした場合、民間の施設や小規模施設についても参加可能性を確保することは重要な事項である。例として一部の大学病院で必須事項として実施している検査が、小規模施設でまったく行われていなかった場合には、必須の入力事項として設定することは困難である。

また海外の施設で通常行われている検討事項が、日本の多くの施設で行われていないようなケースもあり、国内外の臨床体制の違いも注意すべき事項である。また症例の登録基準についても配慮が必要である。ACS NSQIP では症例数が多い大規模施設では、詳細な臨床情報を登録するケースは、無作為抽出により選定された一部の症例に限定する、という基準で入力を行っている⁹⁾。

このように入力項目や入力体制が、各参加施設において現実的かどうかを検証することは必要不可欠な検討事項である。

4. 資源の利用

1) 個人識別コード(unique patient identifier)の活用

米国の社会保障番号(social security number)のような生涯不変の個人識別コードをデータベースに組み込むことができる場合には、異なるデータベース間でも、高い精度でデータを接合することが可能になる。このような個人識別コードの導入は研究面で、臨床データベースの可能性を広げるだけでなく、異なるデータベースで役割分担を行うことにより、コストを軽減することにもつながる。

米国では個人識別コードを用いて、National Death Index と臨床データベースを接合することにより、生存に関する患者フォローアップデータをより正確な精度で得るという研究も多く行われている。また施設間を跨いだデータの連動が可能であれば、患者側も自身の情報を抽出して、自

分自身の診療や予防に活用することができる。したがって個人識別コードを臨床データベースで把握することにより、たとえカルテの保存期間が超過した20年前の手術であったとしても、新しい施設が患者の同意の下に情報を取り寄せ、臨床判断に役立てることができる可能性が生じる。

2) 病院情報システムの活用

各施設の診療情報システムの規格が統一され、かつ電子化された場合には、データベース事業におけるデータ集積のコストを軽減することができる可能性がある。詳細な臨床情報をプログラムにより抽出し、臨床データベースに組み込むことができれば、専門家の判断が必要な一部の項目を除き、各参加施設の入力の手間を省くことに結びつく。診療情報システムの発展と歩調を合わせ、連動した形で臨床データベースの発展させることは有用であろう。

III. 正当性基準 Propriety Standard

1. 基本的人権の尊重、合意形成

1) 研究の倫理指針

多くの臨床データベースは観察研究の枠組みに基づいて行われる事業であり、患者は事業のために検査の追加や、手術、入院期間の延長などはなく、データベースへの登録は診療自体に影響を与えるものではない。このように現実起こった出来事を記録する介入を伴わないプロジェクトについては、現在の日本では文部科学省・厚生労働省策定の「疫学研究に関する倫理指針」⁶⁾に基づいて行う必要がある。

一方で近年では悉皆登録と連動させて、ランダム化比較試験などの介入研究を行うタイプの研究もみられるようになってきている⁷⁾。これは母集団を把握することで、介入のサンプリングコストを下げるだけでなく、介入研究の除外基準の妥当性の検証可能性や、真の母集団との比較による一般化可能性を確保する等の有用性によるものである。ただ、たとえ既存のデータベース事業と同じ枠組み(入力インターフェイスや入力担当者、組

織体制など)で研究実施する場合でも、上記のような介入研究の要素が加わる場合は注意が必要である。介入を伴う研究プロジェクトを追加する場合には、「臨床研究に関する倫理指針」⁶⁾に基づいて、新たに倫理審査を行う必要がある。また被験者に対しても介入別に説明を行い同意書の取得を行うことが必要になると考えられる。ただ上記の「疫学研究に関する倫理指針」や、「臨床研究に関する倫理指針」は近年しばしば改定されているため、基準の推移について十分に配慮する必要がある。また観察研究の枠組み内であっても、臨床データベースの登録対象疾患や登録内容などが拡大することにより、臨床データベースの管理・運営側の性質が変化する可能性もある。

このように基準の変化が著しい場合などには、再度倫理審査を申請し、第三者の判断に照らしてプロジェクトの実施体制を検討することも有用な観点である。

2) 情報セキュリティへの対応

IT技術の進歩と普及により、今日では多くの臨床データベースが何らかの形でウェブを通じて管理・運営されている。ローカルアプリケーションを配布し、あらかじめ決められた時期にデータを事務局に送信するという形式から、ウェブブラウザを通じてリアルタイムでデータを送信するなど形式はさまざまである。しかしながら、ウェブを用いることによって享受される利便性は時として、情報漏洩のスピードを加速させるなど、さまざまなリスクとのトレードオフによって成立していることに注意が必要である。

一方で望ましいセキュリティの基準は、技術の進歩によって変化していく可能性があり、ある時点での評価基準が今後も同様に通用するとは限らないことに配慮が必要である。このような状況では、情報システム管理・運営の対応方針を明文化することは有用な対応であると考えられる。公開した方針の遵守に努め、問題が生じた場合にその内容と対応を共有することにより、臨床データベースの管理・運営者側として一定の説明責任を果たすことができると考えられる。

3) 個人情報の取り扱い

臨床データベースにおけるデータ登録においては個人情報の保護に関する法律に準拠する必要がある。個人情報の取り扱いについては、主として、①連結可能非匿名化情報、②連結可能匿名化情報、③連結不可能匿名化情報、という方針があり、各条件別に配慮が必要である。

また臨床データベースの情報は、たとえ匿名化情報であったとしても、ある一定以上の情報をもった関係者が閲覧すれば、個人を特定できる可能性がある。例として、稀少疾患に対する手術情報であった場合は、関係する臨床スタッフや患者会のメンバーは、手術年や患者の年齢などから、個人を特定することは難しくない。

一方で、臨床データベースは患者側の情報に加え、しばしば参加施設や各施設で治療に関わった医療提供者の情報も含むことがある。したがって調整が加わらない原資料そのものが第三者に渡った場合には、さまざまな流用が行われる可能性がある。このようなデータ流出は、治療成績の非公開を前提としてベンチマーキングを行っている臨床データベースにおいては、参加施設からの信頼を大きく損ねる可能性がある。したがってデータの管理・利用体制に十分に配慮することもまた重要な事項である。

米国胸部外科学会では、申請者から分析計画を受け付け、原資料ではなく分析結果を原則的にフィードバックするという形で対応を行っている⁹⁾。また原資料を提供する場合でも、個人や施設が特定不能な形でデータを加工する、担当者と守秘義務契約を締結する、という対応方法も重要な要素である。

2. 透明性・情報開示

1) データ利用の受付と採択

データ管理・利用においては、公正なスタンスで利用を受け付け・採択することが必要となる。とくに一部施設や企業の不利益となる情報を秘匿し、利益となる情報のみを公開することは、利益相反となる可能性が高い。したがって、データ利

用の受付と採択のプロセスについて透明性を確保し、また採択を行う基準や人選についても情報を開示することは重要な事項である。

2) データ分析結果の公開方法

データ分析結果の公開方法についても、一定の基準を設定することが必要である。米国のように pay for performance によって診療報酬加算が発生するような重症度補正を行う場合には、重症度補正の手順や内容について情報を開示することは必須の事項である¹⁰⁾。

一方で、データ利用申請者が、成果を公開することなく内部資料として用いるにとどめるということも、一つの可能性として考えられる。非公表とした理由が情報の秘匿による利益の独占なのか、公表にたる新規性がなかったのかの判断は簡単ではないが、公正性の観点から公開の方法についても一定の基準を設定することも一案である。

3. バランスの調整

1) 治療成績の公表基準の統一

臨床データベースに登録されたデータに基づいて医療施設の情報を公開する場合に、公表する情報の基準をそろえることは重要である。例として治療成績の指標として「死亡率」を選択した場合にも、ある施設が狭義の術中死亡のみを数え上げることがあれば、他の施設では術後30日死亡というより広義の定義で死亡をカウントする可能性もある。また一方で、死亡数が多く発生した期間を除外した期間で治療成績を算出することもある。また同じ粗死亡率であったとしても、重症度が高い患者を対象に治療を行っている施設と、軽症例のみを対象にして治療を行っている施設では、その意味合いは異なる。

このように、公正な条件が確保されていない中で、治療成績を公開することは、参加施設のバランスを欠き、臨床データベースに対する参加施設の不信感にもつながる。患者側だけでなく、参加施設にとっても公正なバランスで情報を共有することは重要な条件である。

2) 参加の公正性

臨床データベースが広く地域における医療の質向上を目的として事業を行う場合には、各医療施設の参加の公正性を確保する必要がある。米国では診療科単位で、年間数十万の参加費を支払い臨床データベースに参加することが可能である一方、日本では大規模施設であっても同額の費用負担は簡単ではない。広く日本のさまざまな施設からの参加を確保するため、日本の心臓外科領域では臨床データベースの運営主体が入力プログラムを開発し、参加施設側のソフトウェア購入費を無料に設定している。

一方で、データ入力の人的コストは、施設の症例数に比例して大きなものとなる。ACS NSQIPにおいては、年間症例数が一定以上の施設においては、無作為抽出した一定数の症例についてのみ詳細なデータを登録することによって、入力に対する人的コストを一定のレベルに調整するというような対応を行っている⁵⁾。

4. ステークホルダー間の関係、利害対立

1) 重症度補正に対する限界への配慮

性別・年齢や入院時搬送などの限られた項目を用いて行う重症度補正では、疾患のリスクの一部しか説明することができず、臨床現場が理解・納得して用いることができないことはすでに広く指摘されている¹¹⁾¹²⁾。

一方で、臨床データベースのデータを用いた場合には、疾患別の詳細な情報を活用したより有用な指標を確立することができるが、それでもさまざまな限界があることには配慮しなければならない。例として測定していないリスク、稀少なリスクは重症度補正に反映されないため、このような要因を有するハイリスク患者は、依然として医療機関から回避される可能性がある。また治療適応の選択を行う医師個人を情報公開のターゲットとした場合に、患者の治療機会に好ましくない影響を与える可能性があるため、施設を単位とした情報公開が望ましいという指摘もある¹³⁾。

このようにベンチマーキングや情報公開の形式

は時として、患者の治療アクセスに影響することがあるため、十分な配慮が必要である。

2) 専門家の自主性、行政・保険者の政策

臨床データベースにおける利害調整において、最も重要な要素が専門家の自主性の尊重と、行政・保険者のコントロールのバランスである。これは先にあげたように臨床データベースの運営主体にも強く関わる問題である。これらステークホルダー間の関係を考える上では、行政・保険者側の医療費抑制と専門家の診療報酬獲得の利害対立という観点が強調されることが多い。しかしながら近年では、医療の質向上を主眼にした上で、そのためどのようなコストを設定するべきか、という視点で両者が協調して医療を考えることが有用であると指摘されている。医療の質向上にむけた、臨床データベースの活用については、米国の保険者の主たる3つの戦略としてあげられているのは情報公開による患者選択、実績に対する支払い、ベンチマーキング参加に対する診療報酬加算である。

3) 情報公開による患者選択(Public Reporting)

米国ではLeapfrog Groupなどが、病院の症例数やプロセス指標の遵守率、リスク調整済み死亡率を公開している¹⁴⁾¹⁵⁾。一方で、ニューヨーク州、ニュージャージー州などは行政が成績を集め、術者の個人名と治療成績を一般に公開している¹⁾。

このような情報公開の利点としては、高い透明性、施設の名を高め患者を集めることができるという点をあげることができる。

一方で欠点としては、医療提供者が重症の患者を回避すること¹⁶⁾¹⁷⁾、再チャンスの可能性が厳しいことなどをあげることができる。

4) 実績に対する支払い(Pay for Performance)

実績に対する支払い(Pay for Performance)において治療成績を利用する場合、リスク調整を行った上で成績が良好な施設に加算を設定することが例として考えられる。米国メディケア・メディケイド・サービス・センター(CMS)では、CABG手術に対してプロセス指標とアウトカム指標を両方用いて上位20の施設に対して加算を設定してい

る¹⁸⁾。実績に対する支払いは医療施設に対する利益が直接的であるため、治療成績の改善に対する高い動機付けを期待できるという利点がある。

一方で公正な実績の指標を選択することや長期的な指標を支払いに反映させることの難しさが欠点として考えられる。

5) ベンチマーキング参加に対する診療報酬加算 (Pay for Participation)

ベンチマーキングは学会などの同業者内での成績の非公開を前提とした事業であり、参加施設に全体傾向と個々の施設の成績をフィードバックし、成績不良の施設には指導を行う。この方針に対する政策としては事業への参加施設に対して加算を設定するということが考えられ、Blue Cross and Blue Shield of Michigan and Blue Care Network (BCBSM)がこの先駆けとなる政策を実施している¹⁹⁾。非公開を前提としたベンチマーキングでは専門家同士の協力や臨床情報のフィードバックにより、個々の施設の成績の改善につなげ

ることが可能となる。

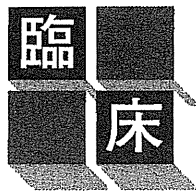
一方で同業者同士の評価による第三者性の低さや、質の低い施設への指導のもつ強制力の低さがデメリットとなる。

先にあげたように、治療成績を指標に用いる上では十分にリスク調整の議論を積み重ねること²⁰⁾、十分な症例数の病院を評価の対象とすること²¹⁾²²⁾、が前提となる。多くの領域でリスク調整が行われておらず、症例数が少ない施設が多数を占めている現時点の日本では、したがって、治療成績をただちに根拠に用いることは難しい。治療成績による評価を行う場合には、ある程度まで集約化を行う必要があることに加え、まずは専門家を中心にしたベンチマーキング事業に Pay for Participation を設定してリスク調整について熟議することが有用であると考えられる。その後に治療成績の「情報公開」や Pay for Performance の設定などが、医療の質向上に向けて現実的に選択可能な政策となると考えられる。

文 献

- 1) New York State Department of Health. Adult Cardiac Surgery in New York State 2004-2006. 2009.
http://www.health.state.ny.us/diseases/cardiovascular/heart_disease/docs/2004-2006_adult_cardiac_surgery.pdf (accessed at 2010/03/09)
- 2) Shahian DM, Normand SL, Torchiana DF, et al: Cardiac surgery report cards: comprehensive review and statistical critique. *Annals of Thoracic Surgery* 72: 2155-2168, 2001.
- 3) Landon BE, Normand SL, Blumenthal D, Daley J: Physician clinical performance assessment: prospects and barriers. *JAMA* 290: 1183-1189, 2003.
- 4) Schneider EC, Spstein AM: Influence of cardiac-surgery performance reports on referral practices and access to care—a survey of cardiovascular specialists. *New England Journal of Medicine* 335: 251-256, 1996.
- 5) Khuri SF, Daley J, Henderson W, Barbour GJ, Lowry P, Irvin G, Gibbs J, Grover F, Hammermeister K, Stremple JF, Aust JB, Demakis J, Deykin D, McDonald G and Participants in the National Veterans Administration Surgical Risk Study: The National Veterans Administration Surgical Risk Study: risk adjustment for the comparative assessment of the quality of surgical care. *Journal of the American College of Surgeons* 180: 519-531, 1995.
- 6) 文部科学省, 厚生労働省. 疫学研究の倫理指針. 2008年12月1日一部改正.
http://www.niph.go.jp/wadai/ekigakurinri/H20_12_01_shishin-all.pdf (accessed at 2010/03/09)
- 7) Serruys PW, Morice MC, Kappetein AP, Colombo A, Holmes DR, Mack MJ, St_hle E, Feldman TE, van den Brand M, Bass EJ, Van Dyck N, Leadley K, Dawkins KD, Mohr FW: SYNTAX Investigators. Percutaneous coronary intervention versus coronary-artery bypass grafting for severe coronary artery disease. *N Engl J Med* 360 (10): 961-792, 2009.
- 8) 厚生労働省, 臨床研究の倫理指針. 2008年7月31日全部改正.
<http://www.mhlw.go.jp/general/seido/kousei/i-kenkyu/rinsho/dl/shishin.pdf> (accessed at 2010/03/09)
- 9) Society of Thoracic Surgeons. Guidelines on Use of STS National Database and Database-derived Information. 10th October 2004.
<http://www.ctsnet.org/file/STSNationalDatabasesGuidelines3FINAL.pdf> (accessed at 2010/03/10)
- 10) Shahian DM, O'Brien SM, Filardo G, Ferraris VA, Haan CK, Rich JB, Normand SL, DeLong ER, Shewan CM, Dokholyan RS, Peterson ED, Edwards FH, Anderson RP: Society of Thoracic Surgeons Quality Measurement Task Force. The Society of Thoracic Surgeons 2008 cardiac surgery risk models: part 1—coronary artery bypass grafting surgery. *Ann Thorac Surg* 88(1 Suppl): S2-22, 2009.
- 11) Khuri SF, Daley J, Henderson W, Hur K, Demakis J, Aust JB, Chong V, Fabri PJ, Gibbs JO, Grover F, Hammermeister K, Irvin G 3rd, McDonald G, Passaro E Jr, Phillips L, Scamman F, Spencer J, Stremple JF: The Department of Veterans Affairs' NSQIP [National VA Surgical Quality

- Improvement Program]: the first national, validated, outcome-based, risk-adjusted, and peer-controlled program for the measurement and enhancement of the quality of surgical care. *Ann Surg* 228 : 491-507, 1998.
- 12) Grover FL, Shroyer AL, Hammermeister K, Edwards FH, Ferguson TB Jr, Dziuban SW Jr, Cleveland JC Jr, Clark RE, McDonald G: A decade's experience with quality improvement in cardiac surgery using the Veterans Affairs and Society of Thoracic Surgeons national databases. *Ann Surg* 234 : 464-474, 2001.
 - 13) Shahian DM, Edwards FH, Ferraris VA, Haan CK, Rich JB, Normand SL, DeLong ER, O'Brien SM, Shewan CM, Dokholyan RS, Peterson ED: Society of Thoracic Surgeons Quality Measurement Task Force. Quality measurement in adult cardiac surgery: part 1-Conceptual framework and measure selection. *Ann Thorac Surg* 83 : S3-12, 2007.
 - 14) The Leapfrog Group fact sheet 2004. (Accessed February 12, 2007)
http://www.leapfroggroup.org/about_us/leapfrogfact-sheet.
 - 15) Birkmeyer JD, Dimick JB: Potential benefits of the new Leapfrog standard: effect of process and outcomes measures. *Surgery* 135 : 569-575, 2004.
 - 16) Burack JH, Impellizzeri P, Homel P, Cunningham JN Jr: Public reporting of surgical mortality. *Ann Thorac Surg* 68 : 1195-1200, 1999.
 - 17) Narins CR, Dozier AM, Ling FS, Zareba W: The influence of public reporting of outcome data on medical decision making by physicians. *Arch Intern Med* 165 (1): 83-87, 2005.
 - 18) Darr K: The Centers for Medicare and Medicaid Services proposal to pay for performance. *Hosp Top* 81 : 30-32, 2003.
 - 19) Birkmeyer NJO, Share D, Campbell DA, Prager RL, Moscucci M, Birkmeyer JD: Partnering with payers to improve surgical quality: the Michigan Plan. *Surgery* 138 : 1815-1820, 2005.
 - 20) Birkmeyer NJO, Birkmeyer JD: Strategies for improving surgical quality-Should payers reward excellence or effort? *N Engl J Med* 358 (8): 864-870, 2006.
 - 21) Peterson ED, Coombs LP, DeLong ER, Haan CK, Ferguson TB: Procedural volume as a marker of quality for CABG surgery. *Jama* 291 (2): 195-201, 2004.
 - 22) Dimick JB, Welch HG, Birkmeyer JD: Surgical mortality as an indicator of hospital quality: the problem with small sample size. *Jama* 292 (7): 847-851, 2004.



臨床データベースにおける科学的質の評価Ⅰ： 医療水準を測定する枠組みの妥当性

*Quality improvement initiative based on National Clinical Database I:
To enhance internal validity and external validity of quality assessment*

宮田 裕章^{*1*3} 大久保 豪^{*2} 友 滝 愛^{*2}

MIYATA Hiroaki

OKUBO Suguru

TOMOTAKI Ai

橋本 英樹^{*1*4} 本村 昇^{*1*5} 村上 新^{*1*6}

HASHIMOTO Hideki

MOTOMURA Noboru

MURAKAMI Arata

後藤 満一^{*1*7} 木内 貴弘^{*1*8} 岩中 督^{*1*9}

GOTO Mitsukazu

KIUCHI Takahiro

IWANAKA Tadashi

臨床データベース(clinical database)は、医療におけるさまざまな分野で臨床現場が主体となって取り組み、発展している事業である。影響を与えるさまざまな立場への有用性や、事業としての実現可能性、法的・倫理的正当性の検証も重要であるが、科学的質の担保も欠くことができない要素である。本稿では根拠に基づいた実践的取り組みとして臨床データベースを位置づけ、医療水準評価を行う上で検証が必要となる妥当性の概念とその手法について検討を行う。

はじめに

臨床データベースは臨床現場だけではなく、患者や一般市民、企業、保険者、政府などさまざまな立場に影響を及ぼす事業であり、管理・運営にあたってはさまざまな観点に配慮する必要がある。著者らの先行研究¹⁾²⁾では、社会的文脈における意義や影響を与えるさまざまな立場への配慮、事業としての継続性や実現可能性、倫理的正当性や関連する法・ガイドラインなど、遵守について検討を行ってきた。これは評価において用いられる有用性基準(Utility standard)、実現可能性基準(Feasibility standard)、正当性基準(Propriety standard)³⁾に基づいた検討である。

一方で臨床データベースの意義を検討する上で、正確性基準(Accuracy standard)に照らし科

学的質を担保することは、根拠に基づいた実践的取り組みとして事業を位置づける上で、欠くことができない要素である。

I. 臨床データベースの科学的質の担保を考えるための視点

科学的質の担保については、著者らが先行研究で整理した概念に基づいて⁴⁾、妥当性 validity (内的妥当性 internal validity)、一般化可能性 generalizability (外的妥当性 external validity)、信頼性 reliability の観点から臨床データベースの科学的質を高めるための手法を検討する。本稿では、この中で妥当性、一般化可能性に焦点を絞り検討を行う。妥当性については一般化可能性に類

^{*1}National Clinical Database ^{*2}東京大学大学院医学系研究科医療品質評価学講座 ^{*3}准教授 ^{*4}東京大学大学院医学系研究科臨床疫学・経済学 教授 ^{*5}東京大学医学部心臓外科 講師 ^{*6}准教授 ^{*7}福島県立医科大学医学部臓器再生外科学講座 教授 ^{*8}東京大学大学院医学系研究科医療コミュニケーション学 教授 ^{*9}東京大学大学院医学系研究科生殖・発達・加齢医学専攻小児医学講座小児外科学・小児腫瘍学分野 教授

Key words : 臨床データベース/医療水準評価/根拠に基づいた医療/妥当性/一般化可能性

する概念も外的妥当性として妥当性に含めるものや、信頼性の概念も含めたより広義な概念として妥当性を位置づけるものもあり、研究によって定義が異なるものであり注意が必要である。本研究では、「知見や仮説の枠組みに自体に対する検証」に関する評価基準を妥当性(内的妥当性)として位置づけ、「知見や仮説の枠組みの外部への適用の検証」に関する評価基準を一般化可能性(一般化可能性)として定義して、検討を行う。

医療水準評価は時系列的な流れとしては、

- ①評価対象とする医療水準の決定
 - ②入力項目・項目定義の明確化
 - ③入力方法・入力体制の構築
 - ④データ整合性の検証
 - ⑤データ分析・結果のフィードバック
- という要素により構成される。

本稿ではこれらの医療水準評価について、妥当性・一般化可能性評価の視座に基づき、

a. 構造・過程・成果などの構成概念が適切に把握されているか

b. 患者視点をはじめ、多角的な視点が反映されているか

c. 医療水準評価が領域のパフォーマンスを包括的に説明しているか

d. 評価に基づいて行動する臨床現場が理解・納得して取り組むことができるか

e. 治療プロセスと成果など、評価に関わる基準の関連は想定された通りか

f. 評価対象となる地域・領域において医療水準評価を適用することが可能か

g. 評価対象外の領域・地域において知見がどのような示唆を与えるか

等の観点から検討を行った。

II. 医療水準を評価する枠組みの構成概念の妥当性

“どのような考え方に基づいて医療水準を把握するのか”という評価の枠組みの選択は、各領域の医療水準の向上を考える上で必須の問いであ

る。先行研究¹⁾でも述べたように、医療の質を考える上では、領域に関わるさまざまな立場に配慮した上で、とくに患者の価値を実現する「品質」を定義・把握し、評価することが重要となる。一方で、どのような領域においても「専門家」という集団が存在する場合には、専門家の活動により、非専門家の活動に比して、異なる効果がもたらされることが前提となる。したがって医療の質を把握する上では、専門領域の定義や活動の内容を把握するとともに、異なる効果を患者の視点から明らかにすることが必要とされる。

医療の質を把握する構成概念としては、構造・過程・結果に注目することが有用である⁵⁾⁶⁾。構造の把握においては、活動を行う組織の構成内容、組織における人的・物的資源の量が必須となる。これに加えて組織の運営・管理の方法、組織文化、組織における情報共有の方法、組織活動に対するインセンティブを明らかにすることも構造を把握する上で有用である。過程とは医療提供者が行う活動の主体であり、治療・手術の内容(投薬の実施状況、手術手技、用いた医療機器など)や診断や検査により構成される。結果は、比較的短期間に観測される死亡や合併症、中長期的な期間で検討する生存率、再発の有無、身体機能などに加えて、主観的な指標である患者満足度やクオリティ・オブ・ライフなどが指標として用いられる。

外科系の専門医制度と連携した National Clinical Database(<http://www.ncd.or.jp/>)では、“一定レベルの手術を適切に実施できる能力を習得し、その臨床応用ができる”という考え方のもと“手術手技に加えて、術前のインフォームドコンセントや周術期管理なども含めて経験する”という基準で、手術経験の構成概念を形成している。したがって外科専門医制度では、同一手術の中で複数の手技が行われた場合でも、重複して数え上げるのではなく、1症例として取り扱う。これは手技単体としての経験ではなく、術前の治療や術後の周術期管理、患者の退院後のフォローも含めて専門家の責任として、取り組みを行うことが背景にある。

治療成績の評価を行う上で、施設や医療提供者別に、患者の重症度を補正してアウトカムを検討することも有用である⁷⁸⁾。この場合、評価対象となる領域のチームの判断で、コントロールが可能な投薬や治療内容を、重症度として扱くと、場合によっては意図的に重症患者が設定されるケースが発生してしまう。したがって被評価者がコントロール容易な項目を除外して、アウトカム指標の重症度補正モデルを構成することが必要とされる。被評価者がコントロール困難な治療開始時の患者の情報を基準としてアウトカム指標の重症度補正モデルを構成することが有用であると考えられる。また分析対象となるアウトカムについても、短期・中長期とは別に、該当する組織の活動の直接的な影響による変化する要素を弁別し、評価対象として構成することが有用である⁹⁾。

III. 医療の質に対する多角的視点の反映

医療水準評価の対象となる枠組みが、単一の視点ではなく、多角的な視点を反映したものとなっていることも有用な条件である。同じ領域の専門家であっても、熟練したスキルの高い専門家だけでなく、経験の浅い修練中の専門家志願者が取り組むことができるようにする必要がある。一方で、比較的大規模な教育研究施設のみで設定した評価枠組みについては、登録に必要となる検査等が小規模施設では行われておらず、参加自体が不可能となることもある。また医療水準評価を領域に関わる専門家で広く共有して、改善に取り組むことも重要であり、この場合、他科の医療チームとの連携を行うことも有用である。外科手術を評価対象とした場合でも、内科や麻酔科医、看護師、MEなど複数の立場の連携により医療水準の評価と改善を検討することは有用である¹⁰⁾。

評価対象となり実施に改善に取り組む臨床現場だけでなく、評価内容が患者や施設責任者、保険者、行政からみても適切かどうかを検証することも不可欠である。体内への医療機器の設置の成功、手術・処置等の手技実施の成功という医療提供者

視点の短期の結果が、生存や予後など患者視点の中長期的な結果にどのように関わっていくかに配慮し、評価枠組みの設計を行うことは不可欠である。一方で改善の取り組みが、他科や施設に、必要以上に負荷をかけるような内容であれば、その実践を継続することは困難であり、診療科・施設としての継続可能かつ、適切なインセンティブが保たれるかどうかを検証することも一案である。また施設の診療成績を優先することにより、救急搬送の拒否や重症患者への治療回避などが、行われることもある¹¹⁾¹²⁾。このような問題に対応する上で、単一の施設診療科だけでなく、領域や地域全体・疾患単位での診療の質を対象にして、検討を行うことも有用である。

また医療政策や保険制度と連動した評価を実施する上では、活動の改善にむけて必要とされる資源や、各種資源の投入の費用対効果、全体として見込まれる経済的効果を、実証的に検証することが有用である。また評価を社会的な文脈に位置づける上では、領域の専門家という内部集団だけでなく、外部の視点を交えて検証結果を示していくことが有用となる。米国 National Quality Forum では、医療水準評価を診療報酬制度と関連づけた提案を行う上で、患者や行政・立法関係者、保険者などの関係者を交えた検討を行い、その成果を公開している¹³⁾¹⁴⁾。

IV. 評価枠組みの包括的説明力

資源の投入、活動、結果、成果とつながっていく一連の活動に対する評価を行う上で、その評価枠組みが医療の質を包括的に説明しているかどうかを検証することが重要である。例として手術実施後の治療内容だけでなく、手術の実施の可否自体の判断も領域の専門医の実力であると考えられる場合には、手術適応の判断に関する項目や、非手術症例も含めて領域の診療の質の検討を行う必要がある。日本でも乳腺外科領域においては、乳癌に対する治療開始を登録開始の基準に設定し、非手術症例も含めて症例を登録し、医療水準評価

を行うことが検討されている。また循環器領域においてはこれまで冠動脈バイパス手術(CABG)実施後の治療成績や、経皮的冠動脈インターベンション(PCI)実施後の治療成績という枠組みで、個別の治療単位で医療水準評価を行ってきた。しかしながら近年のガイドラインではCABGやPCI、積極的薬物療法などの各種の治療を考慮し、状況に応じて最善の治療方針を選択することが領域の専門家により強く求められるようになってきている¹⁵⁾。したがって循環器領域においては今後、治療単位のパフォーマンスのみならず、適応判断も含めた医療水準評価を行うことが有用であると考えられる。

一方で診断や検査など治療過程の一部に限定的に関わり、結果に対しての寄与が相対的に少ない領域を対象にして評価を実施する場合には、結果のみを指標に用いることで領域の水準を説明することは困難である。このような場合には、“エビデンスによって効果が実証されている治療や、領域の専門家によってコンセンサスが得られている治療を、適応のある対象者にどの程度の割合で実施しているか”という、プロセス指標を用いて各診療科の水準の評価を行うことが有益である。

V. 医療水準評価の内容的妥当性

医療水準評価を行ったとしても、評価結果を参照し改善に取り組む臨床現場が、内容について理解・納得できなければ、評価に基づいた改善行動は期待できない。この乖離はとくに、専門家以外の外部集団が評価を主導した場合に発生する現象であり、専門家の主体的な姿勢は継続的な取り組みにおいては不可欠である¹⁶⁾。したがって評価対象とする疾患・治療の構成単位や、指標として設定されるプロセスの内容や適応基準や除外条件、重症度補正治療成績を算出する上での術前リスクの構成などが、臨床現場からみて適切であることは重要な条件である。

内容的妥当性を高める上ではガイドライン設定と同様、RCTをはじめとした先行研究の知見を

参照し、根拠に基づいた枠組みを作成することが重要である。また十分な科学的な根拠が集積していない分野に対しても医療水準の評価を行うことも時に必要とされ、このような場合にはデルファイ変法等を用いたコンセンサス形成により、指標を定めることが一案である。また個々の臨床現場と、改善に向けたコミュニケーションを行うことにより、臨床現場からのフィードバックに基づいて、指標を継続的に改善することも不可欠な要素である。

VI. 評価指標間の基準関連妥当性

医療水準を行う上で測定する指標を個別に把握・検討することは基本であるが、指標間の関連性を検討し、より体系的な視点に基づいて枠組みを検証することは重要である。術前リスクを補正する重症度モデルと治療成績の関連の検討においては、モデルの識別力や適合度を検証することが有用である⁸⁾¹⁷⁾。一方で各種プロセス指標についても、単独での説得力だけでなく、アウトカム指標との関連を検討することにより、介入計画や優先度の設定を行う上で有益であると考えられる。また短期アウトカムと中長期アウトカムの関連を検証することにより、生存患者の予後をより良好なものとするための、一助となる可能性がある。また治療成績の改善による、医療費に対する効果を検証することにより、医療水準評価を位置づけた制度設計を検討することが可能となる¹⁸⁾。

VII. 医療の質評価と改善活動の一般化可能性

活動が行われる領域において、資源の投入→活動→結果→成果という一連の枠組みが一般化可能かどうかを検証することは重要である。ただし、医療水準評価はすべての要素がコントロールされた実験室の中で行われる取り組みではないため、必ずしもすべての領域に対して枠組みが一般化される必要はない。不確実な要素を多くはらむ現実社会での取り組みを検討する上で重要となるの

は、一般化可能な領域と、困難な領域を切り分け、活動の適用可能範囲を把握することである⁴⁾。医療水準評価において、活動の適用範囲に影響を及ぼす要素は、参加する施設・診療科の特徴、患者アクセス・人口動態など地域固有の条件、医療を受給する患者の特徴、保険制度や行政支援などの医療システムの違い、等の影響を考慮することが必要とされる。

著者らの先行研究では、ケースミックスを行った病院標準化死亡比を用いた水準の評価は、総合病院に比して、単科専門病院などの特徴の偏った病院には当てはまりが良好ではないことが示された¹⁹⁾。このように評価の枠組みによっては、評価対象として設定することが困難な領域が発生することがあり、注意が必要である。一方でとくに急性期治療において、患者アクセスの地域差は、潜在変数として影響を与えることが先行研究より示されており²⁰⁾、医療水準の公正な評価を行う上で地域別の救急搬送の状況や、到達前に死亡した事例の把握、救急搬送の受け入れの状況などの把握により対応を行うことも有用である可能性がある。

評価の枠組みの一般化が困難な領域に対しては、知見について外挿可能な要素を判別して情報を提示することが有用である。医療提供体制が異なる地域においては、医療の質改善に向けた適切な政策案は必ずしも同じではない。例として冠動脈バイパス手術においては、米国では基本的に1名の心臓外科医が執刀し physician assistant がサポートを行うのに対し、日本では複数の心臓外科医が手術室に入り手術を行っている。日本には physician assistant に該当する職種はおらず、当然心臓外科医の数も異なる。このような状況で日本の心臓外科医数をただちに米国と同じ割合に

すれば良いかと言えれば必ずしもそれは当てはまらず、少なくとも physician assistant の育成計画も同時に考慮する必要がある。しかしながら必要とされる症例数の差こそあれ、手術に対する一定の経験は治療成績に結びつくことは両国の研究により示されており、この知見は外挿可能なものであると考えられる²¹⁾⁻²³⁾。このように医療水準評価と改善活動の一般化が困難な領域に対しては、条件を吟味した上で、外挿可能な知見、外挿に条件が必要な知見、外挿困難な知見を判別して検討を行うことが有用であると考えられる。

おわりに

臨床データベースの意義を検討する上で、正確性基準(Accuracy standard)に照らし科学的質を担保することは、根拠に基づいた実践的取り組みとして事業を位置づける上で、欠くことができない要素である。医療水準評価の枠組みの適切さを検証する上では、

- a. 構造・過程・成果などの構成概念が適切に把握されているか
 - b. 患者視点をはじめ、多角的な視点が反映されているか
 - c. 医療水準評価が領域のパフォーマンスを包括的に説明しているか
 - d. 評価に基づいて行動する臨床現場が理解・納得して取り組むことができるか
 - e. 評価に関わる基準の関連は想定された通りか
 - f. 評価対象となる地域・領域において医療水準評価を適用することが可能か
 - g. 評価対象外の領域・地域において知見がどのような示唆を与えるか
- 等の観点から検討を行うことが有用である。

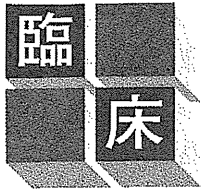
文 献

- 1) 宮田裕章, 後藤満一, 岩中 督, 橋本英樹, 香坂 俊, 本村 昇, 村上新, 木内貴弘, 兼松隆之, 永井良三, 里見 進, 杉原健一, 高本真一: 大規模臨床データベースの意義と展望. 外科治療 5: 332-339, 2010.
- 2) 宮田裕章, 橋本英樹, 本村 昇, 村上新, 木内貴弘, 後藤満一: 臨床データベースの意義と展望 II: 正当性と実現可能性

の検証. 外科治療 6: 797-805, 2010.

- 3) Joint Committee on Educational Evaluation, James R Sanders: The program evaluation standards: how to assess evaluations of educational programs. 2nd edition. Sage Publications, Thousand Oaks, CA. 1994.
- 4) Hiroaki Miyata, Ichiro Kai: Reconsidering evaluation cri-

- teria for scientific adequacy in healthcare research: An integrative framework of quantitative and qualitative criteria. *International Journal of Qualitative Methods* 8(1) : 64-75, 2009.
- 5) Donabedian A : The quality of care - How can it be assessed. *JAMA* 260 : 1743-1748, 1988.
 - 6) Glickman SW, Baggett KA, Krubert CG, Peterson ED, Schulman KA : Promoting quality: the health-care organization from a management perspective. *Int J Qual Health Care* 19(6) : 341-348, 2007.
 - 7) Khuri SF, Daley J, Henderson W, Barbour GJ, Lowry P, Irvin G, Gibbs J, Grover F, Hammermeister K, Stremple JF, Aust JB, Demakis J, Deykin D, McDonald G, Participants in the National Veterans Administration Surgical Risk Study : The National Veterans Administration Surgical Risk Study: risk adjustment for the comparative assessment of the quality of surgical care. *Journal of the American College of Surgeons* 180 : 519-531, 1995.
 - 8) Shahian DM, Blackstone EH, Edwards FH, Grover FL, Grunkemeier GL, Naftel DC, Nashef SAM, Nugent WC, Peterson ED. Cardiac surgery risk models: A position Article. *Annals of thoracic surgery* 78 : 1868-1877, 2004.
 - 9) Hiroaki Miyata, Noboru Motomura, Hiroyuki Tsukihara, Shinichi Takamoto, Japan Cardiovascular surgery database organization : Risk models including high risk cardiovascular procedures : clinical predictors of mortality and morbidity. *The European Journal of Cardio-thoracic Surgery* Nov 1, 2010. (Epub ahead of print)
 - 10) Jacobs JP, Jacobs ML, Mavroudis C, Backer CL, Lacour-Gayet FG, Tchervenkov CI, Franklin RC, Béland MJ, Jenkins KJ, Walters H, Bacha EA, Maruszewski B, Kurosawa H, Clarke DR, Gaynor JW, Spray TL, Stellin G, Ebels T, Krogmann ON, Aiello VD, Colan SD, Weinberg P, Giroud JM, Everett A, Wernovsky G, Elliott MJ, Edwards FH : Nomenclature and databases for the surgical treatment of congenital cardiac disease—an updated primer and an analysis of opportunities for improvement. *Cardiol Young* 2(Suppl) : 38-62, 2008.
 - 11) Burack JH, Impellizzeri P, Homel P, Cunningham JN Jr : Public reporting of surgical mortality. *Ann Thorac Surg* 68 : 1195-1200, 1999.
 - 12) Narins CR, Dozier AM, Ling FS, Zareba W : The influence of public reporting of outcome data on medical decision making by physicians. *Arch Intern Med* 165(1) : 83-87, 2005.
 - 13) Shahian DM, O'Brien SM, Normand SL, Peterson ED, Edwards FH : Association of hospital coronary artery bypass volume with processes of care, mortality, morbidity, and the Society of Thoracic Surgeons composite quality score. *J Thorac Cardiovasc Surg* 139(2) : 273-282, 2010.
 - 14) Welch SJ, Asplin BR, Stone-Griffith S, Davidson SJ, Augustine J, Schuur J : Emergency Department Operational Metrics, Measures and Definitions : Results of the Second Performance Measures and Benchmarking Summit. *Ann Emerg Nov 8, Med* 2010. (Epub ahead of print)
 - 15) Patel MR, Dehmer GJ, Hirshfeld JW, Smith PK, Spertus JA : ACCF/SCAI/STS/AATS/AHA/ASNC 2009 Appropriateness Criteria for Coronary Revascularization : A Report of the American College of Cardiology Foundation Appropriateness Criteria Task Force, Society for Cardiovascular Angiography and Interventions, Society of Thoracic Surgeons, American Association for Thoracic Surgery, American Heart Association, and the American Society of Nuclear Cardiology : Endorsed by the American Society of Echocardiography, the Heart Failure Society of America, and the Society of Cardiovascular Computed Tomography. *Circulation* 119(9) : 1330-1352, 2009. (Epub 2009 Jan 8) Erratum in : *Circulation*. 2009 Apr 21; 119, 15 : e488.
 - 16) Porter ME, Teisberg EO : How physicians can change the future of health care. *JAMA* 297 : 1103-1111, 2007.
 - 17) Noboru Motomura, Hiroaki Miyata, Hiroyuki Tsukihara, Shinichi Takamoto, JACVSD Committee : The Risk Model of Thoracic Aortic Surgery in 4707 Cases from Single Race Nationwide Population, via Web-based Data Entry System : The First Report of 30-day and 30-day Operative Outcome Risk Model on Thoracic Aortic Surgery, *Circulation* 118 : S153-S159, 2008.
 - 18) Birkmeyer NJO, Birkmeyer JD : Strategies for improving surgical quality—Should payers reward excellence or effort? *New England Journal of Medicine* 354(8) : 864-870, 2006.
 - 19) Hiroaki Miyata, Hideki Hashimoto, Hiromasa Horiguchi, Kiyohide Fushimi, Shinya Matsuda : Assessment of hospital performance with a case-mix standardized mortality model using an existing administrative database in Japan. *BMC Health Service and Research* 10 : 130, 2010.
 - 20) Stukel TA, Fisher ES, Wennberg DE, Alter DA, Gottlieb DJ, Vermeulen MJ : Analysis of observational studies in the presence of treatment selection bias: effects of invasive cardiac management on AMI survival using propensity score and instrumental variable methods. *JAMA* 297(3) : 278-285, 2007.
 - 21) Peterson ED, Coombs LP, DeLong ER, Haan CK, Ferguson TB : Procedural volume as a marker of quality for CABG surgery. *JAMA* 291(2) : 195-201, 2004.
 - 22) Hiroaki Miyata, Noboru Motomura, Yuichi Ueda, Hikaru Matsuda, Shinichi Takamoto : Effect of Procedural Volume on Outcome of CABG Surgery in Japan. *Journal of Thoracic and Cardiovascular Surgery* 135 : 1306-1312, 2008.
 - 23) Hiroaki Miyata, Noboru Motomura, Shinichi Takamoto : Enhancing the transferability of region-specific findings: characteristics of Japan's cardiovascular surgery system. *Eur J Cardiothorac Surg* 37 : 249, 2010.



臨床データベースにおける科学的質の評価Ⅱ： 医療水準評価に用いるデータの信頼性と中立性

*Quality improvement initiative based on National Clinical Database II:
To enhance reliability and neutrality of quality assessment*

宮田 裕章^{*1*3} 友 滝 愛^{*2} 大久保 豪^{*2}

MIYATA Hiroaki

TOMOTAKI Ai

OKUBO Suguru

本 村 昇^{*1*4} 村 上 新^{*1*5} 木 内 貴 弘^{*1*6}

MOTOMURA Noboru

MURAKAMI Arata

KIUCHI Takahiro

橋 本 英 樹^{*1*7} 後 藤 満 一^{*1*8} 岩 中 督^{*1*9}

HASHIMOTO Hideki

GOTO Mitsukazu

IWANAKA Tadashi

臨床データベース(clinical database)は、医療におけるさまざまな分野で臨床現場が主体となって取り組み、発展している事業である。影響を与えるさまざまな立場への有用性や、事業としての実現可能性、法的・倫理的正当性の検証も重要であるが、データの科学的質の確保も欠くことができない要素である。本稿では根拠に基づいた実践的取り組みとして臨床データベースを位置づけ、医療水準評価を行う上で検証が必要となる、データの信頼性と中立性の確保について、その概念と手法について検討を行う。

はじめに

臨床データベースは臨床現場だけではなく、患者や一般市民、企業、保険者、政府などさまざまな立場に影響を及ぼす事業であり、管理・運営にあたってはさまざまな観点に配慮する必要がある。著者らの先行研究¹⁾²⁾では、社会的文脈における意義や影響を与えるさまざまな立場への配慮、事業としての継続性や実現可能性、倫理的正当性や関連する法・ガイドラインなど遵守について、検討を行った上で、科学的質の確保を検討する上で医療水準評価の枠組みの妥当性に向けた概念と手法を確認した³⁾。これは評価において用いられる有用性基準(Utility standard)、実現可能性基準(Feasibility standard)、正当性基準

(Propriety standard)、正確性基準(Accuracy standard)⁴⁾に基づいた検討である。

本研究では根拠に基づいた実践的取り組みとして事業を位置づける上で、正確性基準に照らして、妥当性・一般化可能性に加え、信頼性・中立性⁵⁾についてさらなる検討を行う。

I. 臨床データベースの科学的質の確保を 考えるための視点

本研究では、「医療水準評価における観察の安定性の確保」に関する評価基準を信頼性(reliability, dependability)として位置づけ、「医療水準評価におけるデータ提供側とデータ収集側にお

*1一般社団法人 National Clinical Database *2東京大学大学院医学系研究科医療品質評価学講座 *3同 准教授 *4東京大学医学部心臓外科 講師 *5准教授 *6東京大学大学院医学系研究科医療コミュニケーション学 教授 *7東京大学大学院医学系研究科臨床疫学・経済学 教授 *8福島県立医科大学医学部臓器再生外科学講座 教授 *9東京大学大学院医学系研究科生殖・発達・加齢医学専攻小児医学講座小児外科学・小児腫瘍学分野 教授

Key words : 臨床データベース / 医療水準評価 / 根拠に基づいた医療 / 信頼性 / 中立性

る中立性の検証」に関する評価基準を中立性 (neutrality, confirmability, objectivity) と定義して、検討を行う。医療水準評価は時系列的な流れとしては、

- ①評価対象とする医療水準の決定
 - ②入力項目・項目定義の明確化
 - ③入力方法・入力体制の構築
 - ④データ整合性の検証
 - ⑤データ分析・結果のフィードバック
- という要素により構成される。

本稿ではこれらの医療水準評価について、信頼性・中立性評価の視座に基づき、

- ①データ解釈・入力は一貫した経路をたどることができるか
 - ②データ解釈や入力を第三者が検証することが可能であるか
 - ③同じ時期に行われた観察が一致するか
 - ④別の時間に行った観察が一致するか
 - ⑤データ入力側に対してデータ収集側の中立性は保たれているか
 - ⑥データ収集側に対してデータ入力側の中立性は保たれているか
 - ⑦人為的ミスは対策・管理されているか
- 等の観点から臨床データベースにおけるデータの科学的質を確保するための手法を検討した。

II. データ解釈・入力における一貫性の確保

医療水準評価のためのデータ収集における原資料(カルテ, 手術台帳, 検査結果, 看護記録など入力判断の基となるもの)の解釈・データ入力においては, 入力を行う組織, 入力を行う担当者, 入力フォーマット, 入力を行う地域, 入力時期など, さらに医療水準評価の国際間比較を行う場合には, 入力を行う国(地域)によっても異なるものとなることに十分に注意が必要である。したがって, 原資料の解釈やデータの入力に対して, 意図せざる変化が生じる可能性が少なくなるように, 一貫性の確保を行うことが重要となる。

“現在の喫煙の有無”という簡便な項目1つを

とつても, どのような基準をもとに事象の有無を判断したかについて施設や医師により異なる可能性がある。この場合は, 過去に遡って何処までを“現在”とするか, どのような行為を喫煙として定義するのかについて明確にする必要があり, “術前1ヵ月以内に, 喫煙(紙巻き, 葉巻, パイプ, 刻みタバコ等の種類を問わない)”という定義を, 項目とともに明示することは有用である⁴⁾。

診断名を入力する場合でも, ICD-10 のリストから, 系統的に大分類・中分類・小分類を入力する様式に基づく方法と, 自由記述で思い当たる診断を入力させる方法では, 異なる結果をもたらす可能性が高い。前者では “I12.0 腎不全を伴う高血圧性腎疾患”として一貫して分類されたものが, 後者では単に “高血圧性腎疾患” や “腎疾患” などと記述され, 腎不全を伴うかどうかや高血圧性かどうか, など他の分類との区別が困難となる可能性がある。一方で, 基準を統一したとしても思い当たる診断を任意に選択する方法では, 本当にその疾患が存在したかどうかについて, 失念する可能性も高い。とくに記入可能な診断名に上限がある場合には, 治療対象ではない疾患やマイナーな疾患など注目を集めにくい診断名は選択される機会が少なくなる傾向がある⁵⁾。対象とする疾患が存在したかどうか, より漏れが少ない形で把握したい場合には, 手間がかかるが “腎不全を伴う高血圧性腎疾患の有無” を明確な定義とともに, 症例ごとに質問することが有用である。

また地域によって診断に前提とされる常識が異なることにも注意が必要である。冠動脈の3枝病変の定義について, 米国では “冠動脈造影のいずれかの view にて50%以上の狭窄を認める major coronary system の数を記載する。ただし左冠動脈主幹部の50%以上の狭窄は2枝病変に相当する。” という定義を用いているのに対し, 日本では “冠動脈造影のいずれかの view にて75%以上の狭窄を認める major coronary system の数を記載する。” という定義を用いていた。このケースでは個々の病変の定義だけでなく, 病変数の数え上げのルールも異なる。日本では, 日本におけ

る日常診療への配慮とともに、米国との比較可能性を確保するために現時点では、項目を弁別し、2つの基準における入力で整合性の検証を行っている。

入力項目と定義を明確化したとしても、客観的な指標のみで必ずしも定義を構成できるわけではない。とくに、現在の臨床データベースは既存の資料を元にデータを入力するため、治療・手術時から時間が経過するほど、原資料の解釈やデータの入力における、記憶が不鮮明となることは避けられない。また時間が経過しすぎると担当者の異動や、原資料の保存義務期間の超過(<http://www.mhlw.go.jp/shingi/2004/06/s0624-5e.html>)により事後的な判断が不可能となることもある。このような問題に対応する上で、入力可能時期の設定・制限を行うことは有益である。一定期間ごと(1年、6ヵ月)に締め切り時期を設定することも重要である。

一方で、治療開始前・治療中・治療後など個別の情報を、それぞれ直後に入力することもより正確な情報を得る上で有用であるが、厳密な入力時期を設定した施設とそうでない施設では欠損値の発生率に差がないという先行研究⁷⁾もあり、利便性を勘案すると、現時点では必須条件ではない可能性もある。

項目内容と定義を明確に設定し、入力フォーマットを統一したとしても、入力担当者がそれらを十分に理解していない場合には、内容の一貫性は保たれない。項目内容や定義を理解し、原資料を適切に解釈した上で、正しい手順で入力を行う上で、一定の教育プログラムを提供することは有用である。また項目内容や定義の更新や、地域別の状況の変化、注意事項の喚起を行う上で、一定期間ごとに担当者との情報共有を行うことも有用な対策であると考えられる。

III. データ解釈・入力の追跡可能性の保証と検証

さまざまな条件により原資料の解釈、データの入力に差異が生じる可能性がある場合には、その

差異を前提として、原資料の解釈やデータ入力の追跡可能性を保証し、検証作業を行うことは重要である。追跡を行う上では、担当者の明確化、原資料の把握、検証可能な他の情報源の参照などの条件を整えることが重要な条件となり、追跡可能性の保証を行うだけでも一定の効果はある。また一定の手続きの中で、問い合わせや施設訪問による監査を行うことで、整合性を検証するとともに、原資料の解釈や入力における発展的な課題を検討することが可能となる。

原資料の解釈とデータ入力の適切さを検証する上では、入力されたデータと対比する情報源を明確にすることが重要となる。医療水準評価において第1に重要となるのは、入力基準に該当する症例が、漏れることなく登録されているかという登録率である⁸⁾。

例として、死亡が重要な指標となる評価において、死亡症例の登録漏れが発生していた場合には、少数であっても結果を大きくゆがめることにつながる。症例の登録率が検証されない状態で、安易に施設間の治療成績の比較を行うことは、ベンチマーキングを行う上でも間違った結果をフィードバックするリスクとなる。症例の登録率については、手術であれば各施設における手術台帳などを、参照資料として用いることができる。また診療報酬請求において用いられたデータや、他の保険者・行政などへの報告に用いられたデータも対比資料として用いることができる。ただし、登録条件が詳細な場合には、必ずしも精度の高い比較を行うことはできないことに注意が必要である。

登録された症例における、個々の項目の登録内容の整合性の検証においては、その難易度は条件による異なる。入院、手術、生年月日、性別など、どの施設でも必ず記載されており、かつ非専門家でも正誤の判断が容易な項目がある一方で、治療法・術式の詳細や合併症の有無など、標準的な記録が必ずしも行われておらず、また非専門家による判断が困難な項目がある。したがって、個々の項目の登録内容の検証を行う上では、原資料として記録が存在するか、標準化された形式で原資料

表1 データ解釈・入力の見直しにおける区分例

	非専門家による検証が可能	非専門家による検証が困難
(比較的)標準化された形式で 原資料に記載されている。	入院日, 手術日, 生年月日, 性別, 在院死亡, 救急搬送の有無,	診療報酬請求術式, 診療報酬請 求時の主診断, 標準化情報に基 づいた臨床所見
原資料の記載が標準化されてい ない(原資料が残されていない)	30日後状態, 長期生存, 疾患の 既往歴, 検査値	治療後合併症, 標準化されてい ない診断所見, 治療内容の詳細

に記載されているか、非専門家による検証が可能かどうかを第1に検討することが有用である(表1)。原資料における記載が標準化されていない項目については、とくに記載様式の多様性や参照元となる原資料における記載場所を体系的に把握することが有用である⁹⁾。

データの追跡を行う上では、追跡経路を保証することも重要となる。ある期間における登録されたデータの内容に対して責任者を任命し、継続的に連絡体制を構築することは前提となる条件である。医療水準評価においては診療領域別の専門性に配慮した連携体制を設定することが重要となる。ベンチマーキングを行う上で、分析結果のフィードバックを基に、現場の改善に向けた取り組みを促進するためには、臨床現場との連携も視野に入れた連携体制も有用である。ACS NSQIPでは看護師を中心としたデータ入力と整合性検証を行うネットワークと、臨床現場の外科医を中心としたデータの活用と改善を目的としたネットワークが設定されている¹⁰⁾。

上記の資源を用いて実際に、原資料の解釈とデータ入力の過程を追跡し監査を行うことも重要な手続きである。専門家が全施設・全症例に対して監査を行うことは最も精度が高い方法であるが、同時に最もコストが発生する方法であり、膨大なデータの入力が行われる医療水準評価においては現実的ではない。この場合には対象施設、対象症例、対象項目の選定を系統的に行うことにより目的とする整合性検証を行うことが有用である。

また監査は実際に現地施設を訪問しての検証、原資料(複写物や電子データ)の提出による検証、

担当者に対する電話やメールによる問い合わせなど複数の方法がある。非専門家による標準化データを中心とした検討¹¹⁾では、限定された項目のみが対象となり、医療水準評価が対象とする項目をすべてカバーすることはできないが、汎用性の高い方法であり有用である。項目定義や記載方法を改善し、より客観的な手順による監査が実施できるようにすることは重要な方法である。

例として、術後出血合併症についても、単に“72時間以内の出血性合併症の有無”という定義を設定すれば、事後的な検証は困難であるが、“3g/dl以上のヘモグロビン値の低下”と定義すれば、術前術後のヘモグロビン値を検証することで、合併症の有無を非専門家が検討することが可能となる。しかしながら一方で、“3g/dl以上のヘモグロビン値の低下”という条件以外でも出血性合併症が発生している可能性があり、項目の設計には注意が必要である。

IV. 観察の同時的・経時的な一致

データの信頼性を確認する上で、同時期に行われた原資料の解釈やデータ入力が一貫しているかどうかを検討することは、有用な方法である。同じ原資料に対してデータ入力を行った場合に、複数の評価者間で一致した結果が得られるかどうかを確認することや、同じ事象に対して異なる入力経路から報告された結果が一致するかを検討するなどが手法を例としてあげることができる。

一方で観察内容に通時的な安定性があるかどうかを検証することも重要である。領域によっては、診断や検査法、検査結果の判断が基準が通時的に

変化することもある。このような場合には評価者がある時点で行った評価と、別の時点で同じ評価者が同じ原資料を基に適切に行った評価が、異なることが前提となる。このような場合には整合性を検証し、変換や読み替えを行うことにより、ある程度一貫した形でデータを用いることが可能となる¹²⁾。基準が変化する時点を把握し、その影響の検証を行うことは、医療水準評価による経時的影響を把握する上でも必要である。

V. データ入力側・収集側における中立性の確保

医療水準評価においては、データ入力側が原資料の解釈やデータの入力において、意図的に結果をゆがめる可能性も考慮しなければならない。医療水準評価が、施設や医師を非公表としたベンチマーキングとして、あくまでも内部資料として用いられる場合は、この点はそれほど大きな問題とはならない。一方で施設や医師の実名を特定して、個々の治療成績を一般やマスメディアに公表する場合には、とくに中立性への配慮が必要となる。症例単位では、予後が不良であり、登録が不利となる症例情報の除外や、有利な条件となる架空症例の登録が懸念される一方で、個別のデータにおいてもリスクの水増し、有害事象の過小報告などが行われることがある。また登録症例が専門医制度の申請に用いられ、経験症例数や治療成績などで評価を行う場合にも、入力側に入力内容によって異なる利益が発生するため、中立性への配慮を行うことが必要である。

データ入力側の中立性の確保の対策としては、

- ①一定期間(年別)の個別施設の登録データに対する責任者を明示すること
- ②利害が少ない担当者が入力を行うこと
- ③入力プロセスが透明であり、入力内容の相互チェックを行うことができること
- ④原資料と登録データの整合性確認により虚偽のデータが登録されていた場合に、悪質な場合には一定のペナルティが発生するなどの対応を例としてあげることができる。

一方で医療水準評価においては、データ収集側となる運営主体も中立性を損なわないように配慮することが重要となる。とくに運営主体が臨床現場のメンバーによって構成される場合には、入力側との完全な独立性の確保は困難な場合があり、運営主体の構成のみでなく、運営方法についても一定の手順で中立性の確保を検討することが有用となる。評価の際に重症度補正を行った場合に、その手続きがブラックボックスであれば、一部の施設に有利となるような枠組みが意図的に形成された、という疑念は払拭されず、フィードバックの正当性は損なわれる。したがって治療成績のフィードバックを行う上では分析・評価において再現性のある手法を用いた上で、その手順を公表することが重要となる¹³⁾¹⁴⁾。またデータの中立性を確保するためにも、データベースやデータの管理等の運営方法については、一定の第三者性を確保することが重要である。

さらに資金や利害関係についても明らかにし、利益相反の開示や、必要に応じて事実関係の確認や適切な対応方法について検討する必要がある¹⁵⁾。

VI. 意図的ではない誤りへの対応

データ入力側、収集側の中立性を一定のレベルで確保した場合でも、意図的ではない誤りにより、医療水準評価が影響を受ける可能性がある。登録基準の誤認による過小・過剰登録、入力項目や定義の誤認による間違った値の入力、分析手順や入力プログラムの間違いによる間違ったフィードバック、などを例としてあげることができる。一方でこれら意図的ではない人為的ミスが、一定の確率でランダムに発生することは避けることができない現象である。

意図せざる人為的ミスへの対策としては、一定の割合でミスが発生することを前提として、注意喚起やチェックを継続的に実施することが基本となる。医療水準評価における具体的な対策としては、登録対象症例の入力率の通知・確認、個々の症例における項目の入力状況の通知・確認、不正

な入力値や変数間で論理的に矛盾する入力値の入力を禁止するプログラムの実装, 現実的ではない入力値, 変数間で極端に発生頻度が低いと思われる組み合わせに対する注意の喚起, システム利用者の反応を継続的に確認する, などが有用であると考えられる。

おわりに

臨床データベースの意義を検討する上で, 正確性基準(Accuracy standard)に照らし科学的質を確保することは, 根拠に基づいた実践的取り組みとして事業を位置づける上で, 欠くことができない要素である。医療水準評価におけるデータ解釈・入力の信頼性を高める上では,

①原資料の解釈・データ入力に一貫性が確保されるように検討を行うこと,

②データ解釈や入力における追跡可能性を保証し, 必要に応じて監査を行うこと,

③同じ時期に行われた観察の一致や, 別の時間に行った観察の一致を検証すること, が有用である。

一方でデータおよびフィードバックの中立性を確保する上では,

①データ入力側に対する影響への配慮と対策,

②データ収集側に対する影響への配慮と対策を行うとともに,

③意図せざる誤りに対しても対策・管理を行うことが重要である。

文 献

- 1) 宮田裕章, 後藤満一, 岩中 督, 橋本英樹, 香坂 俊, 本村 昇, 村上 新, 木内貴弘, 兼松隆之, 永井良三, 里見 進, 杉原健一, 高本眞一: 大規模臨床データベースの意義と展望. 外科治療 102: 332-339, 2010.
- 2) 宮田裕章, 橋本英樹, 本村 昇, 村上 新, 木内貴弘, 後藤満一: 臨床データベースの意義と展望 II: 正当性と実現可能性の検証. 外科治療 102: 797-805, 2010.
- 3) 宮田裕章, 友滝 愛, 大久保豪, 本村 昇, 村上 新, 木内貴弘, 橋本英樹, 後藤満一, 岩中 督: 臨床データベースにおける科学的質の評価 II: 医療水準評価に用いるデータの信頼性と中立性. 104: 381-386, 2011.
- 4) Joint Committee on Educational Evaluation, James R Sanders: The program evaluation standards: how to assess evaluations of educational programs. 2nd edition. Sage Publications, Thousand Oaks, CA, 1994.
- 5) Miyata H, Kai I: Reconsidering evaluation criteria for scientific adequacy in healthcare research: An integrative framework of quantitative and qualitative criteria. International Journal of Qualitative Methods 1: 64-75, 2009.
- 6) Miyata H, Hashimoto H, Horiguchi H, Fushimi K, Matsuda S: Assessment of hospital performance with a case-mix standardized mortality model using an existing administrative database in Japan. BMC Health Service and Research 10: 130, 2010.
- 7) Clarke DR, Breen LS, Jacobs ML, Franklin RC, Tobota Z, Maruszewski B, Jacobs JP: Verification of data in congenital cardiac surgery. Cardiol Young 18(Suppl 2): 177-187, 2008.
- 8) Halm EA, Lee C, Chassin MR: Is volume related to outcome in health care? A systematic review and methodologic critique of the literature. Ann Intern Med 137(6): 511-520, 2002.
- 9) Society of Thoracic surgeons Adult Cardiac Surgery Database Training Manual
<http://www.sts.org/sections/stsnationaldatabase/datamanagers/congenitalheartsurgerydb/taskforce/articles/trainingmanual2.61.html> accessed at 2010/11/24
- 10) Khuri SF, Daley J, Henderson W, Hur K, Demakis J, Aust JB, Chong V, Fabri PJ, Gibbs JO, Grover F, Hammermeister K, Irvin G, McDonald G, Passaro Jr, E Phillips, L Scamman, F, Spencer J, Stremple JF, the Participants in the National VA Surgical Quality Improvement Program: The Department of Veterans Affairs' NSQIP. The first national, validated, outcome-based, risk-adjusted, and peer-controlled program for the measurement and enhancement of the quality of surgical care. Annals of Surgery 228: 491-507, 1998.
- 11) Maruszewski B, Lacour-Gayet F, Monro JL, Keogh BE, Tobota Z, Kansy A: An attempt at data verification in the EACTS Congenital Database. Eur J Cardiothorac Surg 28(3): 400-444, 2005. (discussion 405-406)
- 12) Quan H, Li B, Saunders LD, Parsons GA, Nilsson CI, Alibhai A, Ghali WA: IMECCHI Investigators. Assessing validity of ICD-9-CM and ICD-10 administrative data in recording clinical conditions in a unique dually coded database. Health Serv Res 43(4): 1424-1441, 2008.
- 13) Khuri SF, Daley J, Henderson W, Barbour GJ, Lowry P, Irvin G, Gibbs J, Grover F, Hammermeister K, Stremple JF, Aust JB, Demakis J, Deykin D, McDonald G, Participants in the National Veterans Administration Surgical Risk Study. The National Veterans Administration Surgical Risk Study: risk adjustment for the comparative assessment of the quality of surgical care. Journal of the American College of Surgeons 180: 519-531, 1995.
- 14) Shahian DM, Blackstone EH, Edwards FH, Grover FL, Grunkemeier GL, Naftel DC, Nashef SAM, Nugent WC, Peterson ED: Cardiac surgery risk models: A position Article. Annals of thoracic surgery 78: 1868-1877, 2004.
- 15) Weinfurt KP, Hall MA, King NM, Friedman JY, Schulman KA, Sugarman J: Disclosure of financial relationships to participants in clinical research. N Engl J Med 361(9): 916-921, 2009.

XI. 特 論

多施設共同大規模データベースの意義

Significance of the national clinical database

岩中 督^{1,3} 宮田裕章^{2,3} 大久保 豪^{2,3}

Key words : 多施設共同, データベース, 専門医制度, 医療水準

はじめに

外科医不足, 医師の地域偏在, 病院医療・救急医療の崩壊などが叫ばれて久しい。病院勤務の外科医は劣悪な職場環境の中で夜間休日を問わず働いている。しかしいくら外科医が頑張っても, その結果が他の地域や他施設と客観的に比較検討されねば社会からきちんと評価されない。同時に外科医は自らが実施した治療の結果を集積し分析することで, 初めて自分たちの次なる目標を設定し, 更なる改善に取り組んでいこうとするインセンティブを働かせる。大規模臨床データベースは, 臨床現場との連携により体系的なデータ集積と実証的な分析を行う基盤事業と位置づけられており¹⁾, より良い医療を長期的に提供することができる体制を構築するためには, できるだけ大規模なデータベースを準備するに越したことはない。そしてこのデータベースで集積されたデータに基づいて様々な課題に取り組み, それらの改善策を検討することで, 個々の専門領域の専門医たちは自らの専門性を社会に示すことが可能となる。

外科関連専門医制度委員会は, 外科専門医制度の基盤となる日本外科学会と各サブスペシャリティの学会で構成された委員会であるが, 外科医が関与している外科手術を体系的に把握し,

更なる外科医療の発展を目的として, 外科専門医制度と連携した症例登録のデータベース事業を2011年度より開始することとした^{2,3)}。

本稿では, このデータベース事業についての背景, 組織の構築の経過, 組織の全容, その事業内容とともに, このデータベースに代表される多施設共同大規模データベースの意義などについて概説する。

1. データベース事業の背景

a. 臨床データベースの必要性

日本外科学会ならびに専門医制度上サブスペシャリティに位置づけられている各外科領域学会で結成する外科関連専門医制度委員会では, 各外科領域の専門医の適正配置・適正数などを把握しておく必要があると判断し, 2008年頃より各学会・団体が集計した調査結果を分析し, 最終的には外科系全体の臨床データベースを構築していく方向で作業を進めてきた。減少しつつある外科医の実態把握を行う必要性, 外科医の増加をめざす方策などについて意見交換を行い, 外科医が本当に不足しているのか, どの地域でどのような外科医が必要なのか, などの外科専門医の適正配置などのデータを自らが有していないことも自覚していた。そのような状況の下で, 外科系全体の手術症例データベースを

¹Tadashi Iwanaka: Department of Pediatric Surgery, Postgraduate School of Medicine, University of Tokyo 東京大学大学院医学系研究科 小児外科 ²Hiroaki Miyata, Suguru Ohkubo: Department of Healthcare Quality Assessment 同 医療品質評価学講座 ³Tadashi Iwanaka, Hiroaki Miyata, Suguru Ohkubo: National Clinical Database 一般社団法人 National Clinical Database

構築すれば、これらの問題点を解決することができるばかりか、施設・地域の医療水準評価を行うこともでき、その結果をもとに国民に質の高い医療を提供することも可能であると判断した。

一方で当時、心臓血管外科専門医認定機構は、既に認定施設を中心として各症例ごとの risk-adjust された極めて詳細な臨床情報を集計したデータベースを稼働させており、その結果の分析・評価により、①心臓血管外科手術実施状況とその結果の把握、②疾患領域別の医療水準評価、③より良質な医療を市民に提供するための臨床情報のフィードバック、④正確かつ迅速な臨床研究支援のための基盤整備などを開始していた⁴⁾。また、日本消化器外科学会は、関連する 8 学会・研究会とともにがん登録事業と連携可能なデータベースの構築を行うため、2009 年度より制度設計・登録フォーマットの共通化などの作業を始めていた⁵⁾。この両学会の指導の下に、外科関連専門医制度委員会内に、手術症例の多施設共同大規模データベースを構築するためのワーキンググループが設置されたわけである。

b. 専門医制度上の位置づけ

並行して外科関連専門医制度委員会では、2008 年初頭より外科専門医と各サブスペシャリティ学会専門医の申請資格、更新資格の統一・共通化をめざした検討を開始していた。その結果、サブスペシャリティ学会の専門医更新要件が、外科専門医更新要件を包含していると判断されれば、サブスペシャリティ学会の専門医を更新する際に外科専門医を同時に無条件で更新できるようにすることとし、専門医制度の基本的な枠組みを変更することを 2009 年に決定した。この専門医制度の変革を実行に移すためには、各サブスペシャリティ学会と日本外科学会間の各会員情報の共有、施設番号の共有、手術実績としての症例登録データの共有、入力フォーマットの共通化などの作業が不可欠となる。そのため、専門医申請・更新に使用する手術実績情報として、この手術症例データベースに登録される症例を使用することとした。これによ

り外科専門医や各領域の専門医を申請あるいは維持したい日本外科学会会員は、手術を実施するごとに症例入力に義務づけられることになるが、その一方で、同一症例の複数回の様々な臨床情報登録を一度の入力で済ませることが可能になるような入力フォーマットを開発した。外科専門医制度ならびに各サブスペシャリティ学会専門医制度の関係を図 1 に示す。サブスペシャリティ学会は、日本外科学会を基盤とした 2 階建て部分を構築するが、日本消化器外科学会には、関連する 8 学会・研究会による 3 階建て部分を更に包含する形態をとることになった⁶⁾。

2. データベースの構造

上記の目的を達成するために設計したデータベースの構造を図 2 に示す。データベースは以下の 3 階建て構造とした。

(1) 統計的調査：手術時に登録可能な少数の項目により構成される(表 1)。利用者数は約 2 万人(専門医制度にかかわるすべての日本外科学会会員)で、外科手術については全例の登録を前提としており、登録は共通のシステムを用いる。

(2) 医療評価調査：各領域の医療水準を評価する術前・術中・術後の項目より構成される(心臓外科領域は約 200 項目、消化器外科領域はテスト入力開始時で更に 19 項目)(表 2)。利用者数は 1 万-2 万人前後で、システム数は外科サブスペシャリティ領域の専門医資格の数と同程度となる見込みである。

(3) 臨床研究：各種リサーチクエスションを明らかにするための項目をプロジェクト別に追加する。追加項目数は数項目~数百項を想定しているが、すべての施設が入力義務を負うわけではなく、各領域の合意形成のもとに参加施設や入力対象症例の基準の設定が必要となる。利用者数はプロジェクトの性質により異なる。

3. データベース運営の組織

この事業を実施するため、外科関連専門医制度委員会はその加盟学会の了解のもとで、一般社団法人 National Clinical Database (NCD)