

厚生労働科学研究費補助金（地域医療基盤開発推進研究事業）  
総合研究報告書

臨床指標の算出方法の標準化およびリスク調整手法に関わる検討の総括

|       |       |                                      |           |
|-------|-------|--------------------------------------|-----------|
| 研究代表者 | 伏見清秀  | 国立病院機構本部総合研究センター<br>診療情報分析部          | 診療情報分析部長  |
| 研究分担者 | 尾藤誠司  | 国立病院機構東京医療センター<br>臨床研究センター 政策医療企画研究部 | 臨床疫学研究室室長 |
|       | 岡田千春  | 国立病院機構本部総合研究センター<br>臨床研究推進室          | 臨床研究推進室室長 |
|       | 西本裕子  | 国立病院機構本部総合研究センター<br>臨床研究推進室          | 臨床研究専門職   |
|       | 小林美亜  | 千葉大学大学院<br>看護学研究科看護システム管理学           | 准教授       |
|       | 本橋隆子  | 国立病院機構本部総合研究センター<br>診療情報分析部          | 主任研究員     |
|       | 堀口裕正  | 国立病院機構本部総合研究センター<br>診療情報分析部          | 主席研究員     |
| 研究協力者 | 井高貴之  | 国立病院機構本部総合研究センター<br>診療情報分析部          | 主任研究員     |
|       | 下田俊二  | 国立病院機構本部総合研究センター<br>診療情報分析部          | システム開発専門職 |
|       | 川島直美  | 国立病院機構本部総合研究センター<br>診療情報分析部          | システム開発専門職 |
|       | 中寺昌也  | 国立病院機構本部総合研究センター<br>診療情報分析部          | システム開発専門職 |
|       | 新森加奈子 | 国立病院機構東埼玉病院<br>内科・総合診療科              | 医師        |
|       | 山田康弘  | 国立病院機構東京医療センター<br>総合内科               | 医師        |

## 研究要旨

諸外国では、臨床指標算出方法の標準化に向け、管理運営データの二次利用により算出した臨床指標の妥当性検証やリスク調整手法の開発が行われている。しかし、我が国では、これらの臨床指標の妥当性検証やリスク調整は十分に検討されていない。本研究では、臨床指標の算出方法の検証、診療録（カルテ）調査における臨床評価指標の妥当性検証、臨床指標の一般化の検証を行い、それに基づき算出方法の標準化を図ること、アウトカム指標に関するリスク調整手法を検討すること、病院情報システムに格納されている患者個票データを用いた交絡因子の調整を検討することを目的とした。

1年目には、国立病院機構臨床評価指標計測マニュアルから算出された70指標の中から19指標を選定し、平成23年のデータを用いて算出方法の見直しを行った。また、アウトカム指標のリスク調整手法は、予測死亡患者数を選択し、国立病院機構の45病院のDPC対象病院を対象とし、平成23年4月1日～平成24年3月31日のデータを用いて検討した。さらに、病院業務で日常的に発生するデータが格納されている病院情報システムのデータベースから患者個票単位でデータを抽出し、DPCデータと連結させることで、リスク調整や交絡変数での調整を行った。2年目には、臨床指標の妥当性を検討するために、診療録（カルテ）調査を行い、カルテ調査で把握された結果をゴールドスタンダードとし、DPCデータで算出した結果の感度、特異度、陽性的中率、陰性的中率を算出した。また、臨床指標の一般化の検証を行うために、国立病院機構臨床評価指標計測マニュアル（2013年改訂版）と医療の質指標ポータルサイトの共通指標定義プール(PDF)の定義に基づいて、それぞれの分母該当症例数、分子該当症例数、施行率（開始率、処方率など）の平均値・標準偏差・中央値・25%タイル、75%タイルを算出し、算出結果に有意な差があるかを調べた。

算出方法の検証結果では、DPCデータやレセプトデータを利用して作成された各臨床指標の施行率には、5つの共通問題があることが明らかとなった。1)死亡退院患者の影響、2)初回入院患者の同定、3)入院時の持参薬の影響、4)逆紹介患者の影響、5)代替検査の影響であった。診療録調査における妥当性検証の結果では、DPCデータによる分母・分子の算出精度は高く、臨床指標の妥当性は高いことが示唆された。一方、DPCデータによるアウトカム指標の算出精度は、過大評価・過小評価になっていることが明らかとなった。臨床指標の一般化の検証結果では、臨床評価指標と共通指標の算出条件に基づいて算出した結果（施行率・開始率・処方率）には有意に異なっていた。アウトカム指標に関するリスク調整手法を検討結果では、予測死亡患者数を算出するリスク調整式において、オリジナルのCharlsonスコアを用いた場合と、英国で新たに見直されたCharlsonスコアを用いた場合では、死亡患者数を予測する精度は変わらなかった。病院情報システムに格納されている患者個票データを用いた交絡因子の調整の検討結果では、病院情報システムから抽出を試みた変数のうち、DPCデータには存在せず、病院情報システム上に存在する交絡変数の存在と、病院情報システムからも抽出が困難な変数が複数存在することが分かった。

今後は、分母や分子の算出条件や算出方法の違いによる算出精度への影響を検証し、臨床的妥当性についてはデルファイ法を用いて、算出条件の統一化を図る必要があると思われる。

## A . 研究目的

諸外国では、臨床指標算出方法の標準化に向け、管理運営データの二次利用により算出した臨床指標の妥当性検証やリスク調整手法の開発が行われている。近年、我が国でも、DPC データを活用した臨床指標算出の試みが始まっている。

しかし、これらの臨床指標の妥当性検証はほとんど行われていない。また、アウトカムの臨床指標（以下、アウトカム指標）は、バイアスを除外するためにリスク調整が必要となるが、この手法も十分に検討されていない。

本研究では、臨床指標の算出方法の検証、診療録（カルテ）調査における臨床評価指標の妥当性検証、臨床指標の一般化の検証を行い、それに基づき算出方法の標準化を図ること、アウトカム指標に関するリスク調整手法を検討すること、病院情報システムに格納されている患者個票データを用いた交絡因子の調整を検討することを目的とした。

## B . 研究方法

本研究は、倫理面への配慮として、国立病院機構本部中央倫理審査委員会に承認を得て実施した。

### 臨床指標の算出方法の検証

分母・分子の抽出条件の再検討や多角的な評価が必要と判断された 19 指標について、算出方法に見直しが必要かどうかの検討を行った。具体的には、これらの指標の分子の該当患者と非該当患者において、患者特性や施行されていた検査や治療の観点から、比較を行い、分子の抽出方法の妥当性を検討した。加えて、分母の抽出条件を変えて、施行率がどのように変化するかを検討した。

### 診療録（カルテ）調査における臨床評価指標の妥当性検証

検証対象指標は、出血性胃・十二指腸潰瘍に対する内視鏡的治療（止血術）の施行

率、人工関節置換術・人工骨頭挿入術における手術部位感染予防の抗菌薬の 3 日以内中止率と術後感染症の発生率、弁置換術における手術部位感染予防のための抗菌薬の 3 日以内中止率と術後感染症の発生率である。対象患者の属性は DPC データから把握し、臨床指標の抽出精度を検証するために、カルテ調査で把握された結果をゴールドスタンダードとし、DPC データで算出した結果の感度、特異度、陽性的中率、陰性的中率を算出した。

### 臨床指標の一般化の検証

国立病院機構 53 病院（DPC 対象病院）に 2012 年 4 月 1 日から 2013 年 3 月 31 日に入院・退院した患者データを用いて、急性脳梗塞患者に対する早期リハビリテーション開始率、急性心筋梗塞患者に対する退院時アスピリンあるいは硫酸クロピドグレル処方率、乳がん（ステージ）の患者に対する乳房温存手術の施行率、急性脳梗塞患者に対するアスピリン、オザグレル、アルガドロパン、ヘパリンの投与率、急性心筋梗塞患者に対する退院時のスタチンの処方率、大腿骨近位部骨折患者に対する早期リハビリテーション（術後 4 日以内）の施行率、急性胆嚢炎患者に対する入院 2 日以内の超音波検査の施行率、気管支喘息患者に対する吸入ステロイド剤の投与率の 8 指標について、国立病院機構臨床評価指標計測マニュアル（2013 年改訂版）と医療の質指標ポータルサイトの共通指標定義プール(PDF)の定義に基づいて、それぞれの分母該当症例数、分子該当症例数、施行率（開始率、処方率など）の平均値・標準偏差・中央値・25%タイル、75%タイルを算出し、算出結果に有意な差があるかを調べた。

### アウトカム指標に関するリスク調整手法の検討

アウトカム指標のリスク調整手法は、国立病院機構の 45 病院の DPC 対象病院を対象

とし、平成 23 年 4 月 1 日～平成 24 年 3 月 31 日のデータを用いて検討した。まず、リスク調整式は、平成 22 年度のリスク調整式と同様の変数を投入し、平成 23 年度のリスク調整式の係数の再算出を行った。そして、そのリスク調整式から予測死亡患者数を算出した。

平成 22 年度は、入院時の依存症を 1987 年に米国で公表されたオリジナルの Charlson スコアによって重症度の補正を行っている。しかし、近年、医学や医療技術の進歩によって Charlson スコアの見直しが行われている。

そこで、本研究においても、平成 22 年度のリスク調整式の変数の中で用いたオリジナルの Charlson スコアをその見直された Charlson スコアに置き変えて、予測死亡患者数を算出した。死亡患者数の予測精度については、ROC 曲線下の面積 (Area under the curve, AUC) により評価を行った。また、それぞれのリスク調整式から算出した予測死亡患者数から、標準化死亡比の算出も行った。

### **病院情報システムに格納されている患者個票データを用いた交絡因子の調整の検討**

病院業務で日常的に発生するデータが格納されている病院情報システムのデータベースから患者個票単位でデータを抽出し、DPC データと連結させることで、リスク調整や交絡変数での調整を行った。

## **C. 研究結果**

### **臨床指標の算出方法の検討の結果**

検査や治療が適切に行われているかを評価する指標の場合、年齢別や術式別の施行率を比較することで、本来行われるべき患者群に対して適切な医療が行われているかを評価することが可能となり、臨床の実態をできるだけ適切に反映させた結果となった。

脳卒中に関する指標を精査したところ、死亡退院患者が他の指標と比較して、多く含まれて

いる傾向にあった。そこで、死亡退院患者をハイリスク症例や標準化治療が困難な症例として、分母から除外することで、施行率は 2～3% の変化が認められた。

本来の治療対象者を同定するためには発症から入院までの日数も考慮する必要がある。例えば、開頭による外科治療あるいは血管内治療を施行されなかった症例群の発症から入院までの日数は、3 か月から数年経過していた。このため、定期検査目的の入院などの患者が多く含まれていることにより施行率が低下していた。同様に、在院日数が極めて短い患者も同じような傾向が認められていた。

入院患者における薬剤の処方率を評価する指標では、入院時に処方となされていない患者群では、予定入院や在院日数が極めて短い患者が多く含まれていた。

退院後の外来における検査の施行率を評価している指標では、検査が行われていなかった患者群には逆紹介患者が多く含まれていた。これにより、退院後の検査の施行率は低下していた。

### **診療録 (カルテ) 調査における臨床評価指標の妥当性検証の結果**

出血性胃・十二指腸潰瘍に対する内視鏡的治療 (止血術) の施行率の DPC データの分母抽出の陽性的中率、陰性的中率を調べた。その結果、陽性的中率は、A 病院では 62.5%、B 病院では 88.9% であり、陰性的中率は、A 病院では 97.6%、B 病院では 100.0% であった。また、DPC データの分子抽出の陽性的中率、陰性的中率を調べた。陽性的中率は、A 病院では 100%、B 病院でも 100% であり、陰性的中率は、A 病院では 97.5%、B 病院では 97.1% であった。

次に、DPC データにおける人工関節置換術・人工骨頭挿入術術後 3 日以内の抗菌薬中止率に関する陽性的中率、陰性的中率の結果は、陽性的中率は、100.0% であり、陰性的中率も、100.0% であった。また、DPC データにおける感染症の発生率の陽性的中率、陰性的中率は、陽性的中率は、33.3%

であり、陰性的中率は、95.6%であった。

次に、DPC データにおける、弁置換術における手術部位感染予防のための抗菌薬の3日以内中止の算出結果は、カルテ調査の結果とすべて一致した。

### 臨床指標の一般化の検証の結果

急性脳梗塞患者に対する早期リハビリテーション開始率

急性心筋梗塞患者に対する退院時アスピリンあるいは硫酸クロピドグレル処方率

乳がん（ステージ ）の患者に対する乳房温存手術の施行率

急性脳梗塞患者に対するアスピリン、オザグレル、アルガドロパン、ヘパリンの投与率

急性心筋梗塞患者に対する退院時のスタチンの処方率

大腿骨近位部骨折患者に対する早期リハビリテーション(術後4日以内)の施行率、

急性胆嚢炎患者に対する入院2日以内の超音波検査の施行率

気管支喘息患者に対する吸入ステロイド剤の投与率

| 上記の<br>指標番号 | 施行率    |       | 有意差 |
|-------------|--------|-------|-----|
|             | NHO 指標 | 共通指標  |     |
|             | 65.2%  | 45.0% |     |
|             | 93.1%  | 84.1% |     |
|             | 76.0%  | 76.0% |     |
|             | 83.9%  | 55.8% |     |
|             | 93.0%  | 71.5% |     |
|             | 70.1%  | 68.7% |     |
|             | 34.1%  | 34.3% |     |
|             | 70.4%  | 50.4% |     |

乳がん患者に対する乳房温存手術の施行率と急性胆嚢炎患者に対する入院2日以内の超音波検査の施行率は、両指標の算出結果に有意な差は認められなかった。残りの6指標については、両指標の算出結果に有意な差を認めた。

### アウトカム指標リスク調整手法の結果

オリジナル Charlson スコアを用いたリスク調整では、AUC は、0.858(95%信頼区

間：0.855-0.860)であり、予測能の精度は中等度であった。新 Charlson スコアを用いたリスク調整の AUC は、0.857(95%信頼区間：0.855-0.860)であり、予測能の精度も中等度であった。

### 病院情報システムのデータベースを用いた交絡因子の調整の結果

病院情報システムから抽出を試みた変数のうち、DPC データには存在せず、病院情報システム上に存在する交絡変数の存在と、病院情報システムからも抽出が困難な変数が複数存在することが分かった。前者においては、検査部門システムや定型的に記載される看護記録の有用性を確認することができた。後者については、併存症や生活歴、初療時及び入院時にテキストデータとして電子カルテ上に記録されているデータであった。

## D. 考察

### 臨床指標の算出方法の検討の考察

臨床指標の妥当性検証を行う場合、重要な2つの視点がある。一つは、データから抽出された Target population の精度や臨床の実際の状況が反映されているかどうかである。

本研究は、データから抽出された Target population の精度を把握した。その結果、DPC データやレセプトデータを利用して作成された各臨床指標には、5つの共通問題があることが明らかとなった。

一つ目は、死亡退院患者を分母に含めるべきか、除外するべきかという問題である。評価しようとしている指標の治療や検査を行わなかったから死亡したのか、死亡しやすいハイリスクな患者であったのかは、DPC データやレセプトデータからはわからない。よって、死亡の因果関係のわからないものに関しては、安易に解析対象から除外することは望ましくないと考える。

二つ目の問題は、検査の目的や実施するタイミングが明確化されていない検査の

Target populationをどのように抽出するかである。検査の施行率を算出する場合、検査の目的が初回の診断目的なのか、経過観察目的なのかによって、分母をのべ患者数でカウントするか、実患者数でカウントするかに影響する。また、現在では計測期間外に実施されていた検査については、拾うことができない。今後は、これらの影響をどのように考慮していくかを検討する必要がある。

三つ目の問題は、入院時の持参薬情報がないことである。例えば、気管支喘息患者で発作入院を繰り返している患者の多くは、すでに吸入ステロイド剤を処方されていることが多く、3日以内の入院では新たに処方されることは、少ないことが推察される。また、予定入院患者のすでに吸入ステロイド剤が処方されている場合、入院時に薬を持参しているため、入院時の処方率が低下する。処方率を評価する指標の場合、このような持参薬の影響は大きいと考える。

四つ目の問題は、逆紹介患者である。急性増悪した場合に入院し、治療が終わったら、かかりつけ医で外来フォローが行われる場合、計測対象病院では入院後の外来で検査や指導を継続的には行われないため、施行率は低下する。一方で、逆紹介患者が多いということは、地域連携がしっかりしている病院や診療科を示していることでもあり、今後は逆紹介している患者については除外するべきと思われる。

五つ目の問題として、検査・治療で選択肢が多いもの施行率の評価である。例えばガイドラインでは、超音波検査が奨励されていても、現在CT・MRIが普及し、総合的に評価ができるようになってきた。また、静脈血栓予防の代わりに積極的なりハビリテーションの介入などで代替できる場合、その一方だけの施行率を評価することの妥当性の問題がある。

### **診療録（カルテ）調査における臨床評価指標の妥当性検証の考察**

出血性胃・十二指腸潰瘍に対する内視鏡的治療（止血術）のDPCデータ算出による

分母と分子の陽性的中率と陰性的中率は高かったが、陰性的中率に比べ、陽性的中率率が低かった。その原因として、コーディングのミス、医師によるカルテ記載の不備が考えられる。また、カルテ記載方法の統一化や改善を図っている病院では、診療情報管理士が行うコーディングの正確さにも影響していると思われる。

次に、DPCデータにおける「人工関節置換術・人工骨頭挿入術における手術部位感染予防のための3日以内の抗菌薬中止」の算出結果とカルテ調査による結果は、すべて一致しており、算出精度が高いことが示された。一方で、本指標の算出条件で検討が必要な点も明らかとなった。内服薬による4日目以降の予防的投与や術前から抗菌薬を投与している症例などの検討が必要と思われる。

また、DPCデータから術後感染症を同定する感度および陽性的中率は低く、過少評価となることが明らかとなった。今後は、抗菌薬の連続投与日数に関わる条件だけではなく、抗菌薬の再開の有無、再開後の投与期間、投与期間中の抗菌薬の種類変更に関する要件などを加えることにより、精度を高めていくことが求められる。

次に、DPCデータにおける、「弁置換術における手術部位感染予防のための抗菌薬の3日以内中止」の算出結果とカルテ調査による結果は、すべて一致しており、算出精度が高いことが示された。また、病院内での投与期間のばらつきはほとんどないが、病院間での投与期間の違いが明らかになった。弁置換術は清潔手術となるため、3日以内の抗菌薬の中止が基準となるが、ASHPによる心臓胸部外科手術のガイドラインでは、抗菌薬は術後72時間継続を推奨している。今後は、術式別の投与期間についての検討や臨床的な妥当性の検討が必要と思われる。

### **臨床指標の一般化の検証の考察**

6指標の算出結果に有意な差を認められた理由として、分母と分子の算出条件や算出方法の違いが考えられる。

分母の算出条件における問題点としては、各指標が対象とする傷病名の相違、対象症例や除外症例の臨床的妥当性などがある。分子の算出条件における問題点としては、退院時処方と同定方法の相違、薬剤の抽出方法と対象薬剤の種類や数の相違、リハビリテーションの開始時期の臨床的妥当性などがある。

### **アウトカム指標リスク調整手法の考察**

予測死亡患者数を算出したリスク調整式に用いたオリジナル Charlson スコアは、開発されてから 25 年以上経過している。この間、Charlson スコアを算出するための併存症に関する治療方法も進歩し、重症度も変化している。このため、重症度の見直しに基づいて、新たにスコアリングされた新 Charlson スコアを用いて、予測死亡患者数を算出することを試みた。

しかし、Charlson スコア以外の独立変数は変えずに、オリジナル Charlson スコアを用いた場合と新 Charlson スコアを用いた場合でのリスク調整式で、予測精度は変わらなかった。分析対象とした 53 施設のオリジナル Charlson スコアを用いた予測死亡患者数と観察死亡患者数の比である標準化死亡の範囲は 0.53-1.63 であり、新 Charlson スコアは 0.55-1.58 とばらつきにも影響がみられなかった。

今後は、日本人の患者属性を考慮した Charlson スコアを検討していく必要がある。また、今回の副傷病の重症度は、入院時依存症から算出をしたが、入院後発症疾患名を含めて得点化し、種々のリスク調整式を検討する予定である。

### **病院情報システムのデータベースを用いた交絡因子の調整の考察**

DPC データの弱点は、患者個人が持つ固有の情報や、患者の具体的な転帰も含めたアウトカムの情報が少ないことにある。病院情報システムに格納されたデータ、特に、検査部門データおよび看護師による定量化された観察記録については、リスク調整変

数として利用する価値が高く、妥当性や補完率も高いデータであることが示唆された。

一方、病院情報システムからも抽出が困難なデータの多くは、患者の初療時や入院時にテキストで入力されている項目であった。今後は、入力時の負担をかけずに、記録の質も向上し、さらに二次利用の可能性を広げるような入力フォーマットを、電子カルテにアドオンするための工夫が必要であると考えられる。

### **E. 結論**

臨床指標の算出方法の標準化およびリスク調整手法について、現在の DPC データによる分母分子の算出精度やリスク調整式の精度は高く、当該臨床指標の妥当性は高いことが示唆された。一方で、DPC データによる術後感染症の発生率などのアウトカム指標については、過大評価・過小評価になっていることが明らかとなった。また、様々な研究機関や医療機関が臨床指標を作成しているが、その算出結果は有意に異なるものであった。その原因として考えられる分母・分子の除外条件や算出方法を統一化することで、多施設間での比較が可能となる。

### **F. 健康危険情報**

なし

### **G. 研究発表**

1. 論文発表  
なし

2. 学会発表  
なし

### **H. 知的財産の出願・登録状況**

なし