

⑤ 各委員は第2回も個別に評価することで、必ずしも検討会の議論の大勢にとらわれることなく、自らの意見で評価することができる。といった特徴がある。

6. 過程評価の一般的留意点

これら QI の作成過程でいくつかの留意すべき点を以下にまとめた。これらはコンセンサスというわけではないが、筆者がこれまでに行ってきた QI 作成における考え方である。1つの考え方として、これからの検討のきっかけになればと考える。

a. 診療としては行われたが記録されていないものをどう扱うか

標準診療がどの程度行われているのかを吟味するために診療録から情報を収集することから、診療として行われていても、記録がされていない場合をどう取り扱うかという問題がある。これに関しては、記載されていることそのものが標準診療として考えられるかを元に、QI の適切性を吟味した。つまり、診療を行うだけで診療録に残っていても構わないものに関しては、QI としては厳しすぎるということで、不適切と評価するようにパネル委員へは依頼した。そのため、最終的に採択された QI は、記載されていないことは質が低いと考えてよい、という専門家パネルの判断を表している。

b. どのレベルの施設を対象とした QI を作るのか

QI の作成について専門科に相談すると、頻りに上がる質問は「どのレベルの病院を対象とするのか」ということである。おそらくこれは「病院評価」という視点から持ち上がる疑問であろうが、本来医療は患者がどのよ

うな医療を受けているか、という患者単位で評価されるべきものである。そのため、どのレベル、というと「日本の患者であればこの程度の医療水準であるべき」といえる範囲というのが、正確な回答となるのである。患者の視点から見れば、特に大学病院だったら、一般病院だったら、ということから、要求される QI のレベルが変わることはないはずである。

一方で各施設においては担当する診療の種類や範囲は異なるかもしれない。大規模なセンター施設では、長期フォローがなされることは一般病院よりも限られるため、長期フォローを対象とした QI は、大規模センターを評価するには適さない、ということになるかもしれないし、逆に大規模なセンターでは既に診断がついて紹介される場合が多いため、診断過程を問う QI はあまり当てはまらないということになる。これは診療のレベルとは少し別の問題であり、いわばレベル、といった高低の問題と言うよりも、幅の問題であることを意識して考えることが必要になる。

しかしながらこれらの細かい概念はイメージがつかみにくい点もあるため、専門家パネルのプロセスを運用していく上では一定の平均的、最大公約数的な施設を設定することで具体化して QI の検討を進めることがよいかもしれない。

c. 保険の支払いを記載されていない診療を標準としてよいか

本来、標準診療がまず存在し、それらを元に保険による支払いとして記載されるべきであり、保険により標準診療が影響されるべきではない。だから、理想的には QI の設定は保険支払いを無視して考えてよいはずである。しかし、QI は一方で現場の医療に対

してその示す診療が確保されることを要求するといった側面もあり、そうなると、保険収載されていないものを現場に対して要求することになるので酷である。具体的にそれが大きく問題となった QI 候補はなかったが、候補作成の当初より、そういった配慮から今回は保険収載されていないものについて QI とすることは原則として避けるように考えた。

d. 各 QI の重み付けと、総合点計算について

これら、個々の診療行為を規定する QI において、対象患者を分母として定義し、それに対して行われる標準診療を分子として準拠率を計算する構造となっている。準拠率はこれら、それぞれの QI について計算される。それらを見ることで自施設の改善点などが明らかになるが、では、総合点はどうするのか、という問題がある。総合点を計算するためには、それぞれ QI に、重要性を加味した重み付けが必要である。QI の作成時には QI の内容の QI としての適切性は吟味したものの、重要性については特に検討しなかった。また、本来 QI の重要性は、どれだけアウトカムを改善することに寄与するのかということによって決められるべきであるが、治療行為をそのまま対象とした QI ばかりではなく、診断行為の適切性などの QI も多くあり、診断行為は、それに基づく治療行為が適切であって初めてアウトカム改善に結びつくため、それらのアウトカムとの直接的な関連の強さを見るのは困難である。世界的にも QI の総合点の算出方法は未確立なのが現状である。

一方、うまく総合点が算出できたとして、その総合点が何を意味するのか、どう改善に結びつくのかということが逆に不明瞭にな

ってしまう一面もある。個々の QI を吟味していると、改善点はその QI の指定する診療行為にあることは明白であり、それが過程評価の QI の長所でもあった改善点を直接指し示すということを弱めてしまう結果になる。多施設との総合的な「質の比較」をするのが目的であれば総合点の算出は必須であるが、改善という意味においては弱くなるのが皮肉ともいえる。

それでも総合点に関する探求はいくつかある。1 つは QI を作成するときと同様の形で、それぞれの QI の重み付けも専門家パネルの意見を元に設定するというものである。しかし、これはパネルに QI の「妥当性」つまりこれが質の指標としての適切性の強さを問うているのか、期待されるアウトカム改善の大きさを問うているのか、等の概念が混乱する可能性がある。エビデンスは確立しているものの、効果の小さいと考えられるものは、エビデンスはある程度弱点があったとしても効果が非常に大きいと考えられる（少なくともパネルの意見上）ものについては、後者の方が重要性は高いだろうが、QI としての妥当性は、前者と後者の比較において甲乙つけがたい。QI を採択する作業においては、絶対評価であり「甲乙」ということを考えずにすむが、重み付けというと、QI 間の比較といった要素が多分に含まれるためこのような問題が生じてしまうことに留意する必要がある。

また別の QI の総合点を出す方法として、ある分野の QI、例えば糖尿病ならば糖尿病の QI で患者一人一人について、受けた医療が該当する QI を全て満たしているか否かで合否の 2 択としてしまう¹⁹、つまり、患者 A は、当てはまる QI 全てでその通りの医療を受けたとすれば「よい医療」、1 つでも QI 通りでない医療を受けていれば「劣った医療」

を受けたと計算する方法もある。このようなやり方が、米国で広がりつつあるが、これは、例えば、4、5のQIが設定されている場合には、それを全部満たすことは比較的容易である。しかし、20~30ものQIが用意されたとしたら、それを全て満たすのは非常に困難であることは想像に難くない。つまり、「1」と数えられる患者は非常に少ない。ひょっとすると、QIが想定する臨床状況に当てはまらない、つまり「必要な医療が少ない健康な人」だけが1というスコアを得るかもしれない。これを施設平均として何%の患者が1か、という数字を出すと、実は、健康な人々の割合と強く相関するかもしれない。そうすると、先述のProcess-Outcome Linkつまり、「質が高い」医療を受けた人々は「健康になる」という関係が容易に見いだされるかもしれない。もちろんこれは仮説であるが、今後の研究で吟味していく必要がある。もちろん、QIが多くあるときにはいくつかのグループ分けをして、その中でそれぞれ1/0のスコアがあってもよいかもしれない。

e. 実際の測定対象

実測にあたって、作者が考える標準的手法について以下に述べる。まず、対象であるが、基本的に患者リストをまず作成してその患者の診療録を検討することで、QIへの遵守率を調べることができる。その際に、最初から標準診療が行われることがあまり望まれない例については除外して考えることがよいだろう。もちろん、目的によって使い分け、自施設の質を検証するためには、全ての症例を対象として評価した後に理由を考えると、いう方法もありえるが、もし、多施設比較を行って改善の推進力にしようと考えるときなどは、対象症例の範囲を明確にしたデータを収集することが必須になる。以下は、筆者

が思いつくままであるが、QIを適用する際に基本的に除外条件とした考える案である。

・介入を伴う臨床研究に入ったもの

・他院での治療途中で転院してきたもの

このように、比較をする場合には比較する対象を一定の基準でそろえていくことが必要であることを意識する必要がある。

f. 例外の取り扱いについて

患者の希望や併存症により、標準診療が行えないことは往々にしてあるものである。そのような事例をQI適用するにあたってどのように考えるか、というのは大きな問題である。例外を考慮する、ということは通常の診療にあたって必要なことであることは疑いない。しかし、測定を行う際にはその検討には大きな困難が伴う。また、都合の悪い患者を分母から除外するというインチキが行われる可能性もあり、イギリスの国レベルのプライマリ・ケアの質評価・診療報酬プログラム(Quality and Outcome Framework)では例外を報告した割合が質のスコアの強力な予測因子であったという研究結果もある²⁰。しかし、例外をみとめない評価は、盲目的な標準準拠を推進する力として働くためそれは逆に患者にとって不利益となる。ただ、その好ましくない推進力の強さは、単純なフィードバックではそれほど強くないと予想され、逆に診療報酬とリンクさせたシステムの中では強力に働く。また、医療者の評価に対する信頼は例外を考慮することで高まることは予想される。これら諸要素を考慮して、測定方針を決定するべきである。上記イギリスのプログラムは例外の適用をみとめた例であるが、逆にアメリカ外科専門医会・臨床腫瘍学会(American society for Clinical Oncology)・包括的全国がんネットワーク(National Comprehensive Cancer Network)の

作成した QI では例外はみとめずにスコアは計算する方針をとっている²¹。これら、前者が診療報酬リンクプログラムであるのに対し、後者が質改善のためのフィードバックを主眼としていることに起因するものかもしれない。

7. QI を使う情報源について

QI を使用して診療の質を評価するにあたっては、できるだけ労力を少なくして広範囲かつ詳細なデータを収集することが望ましい。しかし、この2つは相反するものであり、詳細なデータを集めるには労力が必要であり、逆に省力では収集データに限りがある。以下にいくつかの情報源の候補とその特徴について考察する。

a. 診療報酬請求情報

診療報酬請求情報は各施設で毎月作成されるものであり、行われた診療行為は請求情報となるため効率的に診療行為情報を収集できる。また、請求にあたって各施設は漏れないように動機付けも存在し、かなり正確な診療行為を抽出することができるのが利点である。しかし、一方で最大の限界は検査結果などの患者の状態が不明であることである。患者の状態や検査の結果がどうであれ診療報酬は変わらないため、検査結果は診療報酬請求情報には表れない。他にも

- ・診療報酬請求上請求できないもの、例えば限度を超えて採取されている腫瘍マーカーの測定などは、請求されないこともありえる。

- ・院外の薬局から処方薬が出される院外処方箋の内容なども診療報酬請求書には掲載されない。

- ・また、数としては少数であるが、自費診療、労災保険の使用などの部分については診

療報酬請求としては含まれてない。

などの限界を認識しつつ、診療報酬請求情報を使用すべきである。

さらに、この診療報酬請求情報を解析するためには電子化されていることが必要であるが、この問題には施設によって様々であることが予想される。電子請求は近いうちに義務化されることが決まっているが、現状としてはまだ限られており、紙ベースの請求が行われている。しかし、実際には手書きのレセプト（診療報酬明細書）というのは稀で、一方で大多数の診療報酬請求書はコンピューターによって作成されて印字されている。そのため、処理に耐える電子データが使用可能かどうかは、その施設のコンピューターが出力可能かどうかによって依存する。

一般に QI を使用するのが施設であるため、診療報酬請求情報といったときには施設からデータを抽出することを考えるのが通例だが、逆に請求される側、つまり保険者の保有する診療報酬請求書を使用するという方法も考えられる。この場合患者単位で受けた医療の請求が集まるため、多施設を同時に受診している患者についても受けた医療行為全ての捕捉が可能である。しかし、内容的には上記のような限界は念頭においておかなければならない。また、保険者の保有する診療報酬請求データは紙ベースであることが多いこと、また査定などによりその部分は別途補完されていたりすることにも注意が必要である。

最近 は 診 断 群 分 類 である Diagnosis-Procedure Combination (DPC) の分類を使って患者の状態を反映させようという試みもなされている。これは、これまでの診療報酬請求情報のみの解析に比較すると格段の進歩であるといえる。しかしながら、こ

の情報はまだ診療報酬目的の利用が主であり、また報酬目的であるが故に研究の枠組みの中でその正確性や妥当性の検証を行うのは壁があるのも事実である。(分類が「正確か」という吟味は「不正の有無」の吟味に直結し、情報の正確性という課題を超えて、立ち入り検査的な色彩を帯びがちである。)今後の検証が待たれる。

また、請求上 DPC は最も医療資源を投入した分類を主病名と考えるのがルールであるため、病状の細かい有無や重症度を反映しているかどうかとも検証を待たねばならない。

b. がん登録

がん診療の分野においては、がん登録がシステムとしては行われており研究等や統計情報として役立てられている。これには学会等が主体となつて行う「臓器がん登録」、がん診療連携拠点病院において義務づけられている「院内がん登録」、都道府県単位で行われている「地域がん登録」がある。臓器がん登録は、治療方法の評価や病期分類の妥当性を検証する目的などのために実施されており、臨床的に詳細な情報が含まれているのが特徴であり、地域がん登録はがん罹患率の推計など統計を目的に作成されている。院内がん登録は両者のいわば中間であり病院が実態を把握して、運営や医療の質の向上を計画していくことを目的として行われている。この中で QI を使用するのに必要な情報が含まれているのは主に臓器がん登録であるが、これは学会が実施主体であることから会員のいる施設からだけデータが収集されること、また診療科が限られている例えば外科症例のみ、内科症例のみといった、会員がいる科のみが提出されるなど、サンプルの範囲が狭い可能性がある。逆に情報は QI を適応するのに適した情報が集まっておりその点で

は QI の適応に適していると考えられる。一方院内がん登録は、網羅性を追求することは完璧とまでは行かないまでも、施設のしくみとして存在するために、内科症例、外科症例などの悉皆性はほぼ確保されているといつてよい。しかし、全がん種で同一の項目を収集していることから、おのずと臨床的な詳細情報の収集は限られており、また、病期分類に日本の取り扱い規約ではなく、UICC の分類を用いていることなどから、QI をそのまま適応することは困難である。しかし、ここから出発してサンプルを確定し、診療報酬請求情報などを収集するための出発としては最適な情報源といえ、また、診療報酬請求情報と組み合わせることによりある程度の QI の適用が可能になる。

がん以外にはあまり疾病登録が行われていないのが、残念であるが、今後疾病登録等進むにつれて、がん以外の分野でも QI の適用が可能になっていくものと考えられる。

c. 診療録採録

できるだけ詳細な臨床状況を把握して総合的な評価を行うためには、診療録採録がどうしても必要になる。医師の判断過程や例外的事項を知る手がかりは、患者、医師に直接尋ねる以外には、診療録以外の情報源は考えられず、貴重な資料である。

ただし、診療録の採録は非常に手間のかかる作業である。診療録の構造も各科で分けられているものから入院・外来だけ分けられているもの、など様々であり情報を収集するに当たって、採録者がその構造に慣れる必要がある。また、悪筆乱字、日本語、英語、ドイツ語がアルファベットとカタカナで入り交じり、さらにスペルが間違っていることも多々あるため、内容を理解できる者が判読しないと難しい。そのため必然的に医療資格を

有する者が採録することが必要になる。

電子カルテについて、可読性は非常によくなるものの、採録作業は依然困難になる。電子的なデータベースとなると個々の患者の範囲が明確でなく、縦横無尽にデータが付加されていくことになる。また、通常1回分、1人分の記載で区切られてしまうため移動しながら全体の流れをつかむのが大変である。また、前のカルテを複写して変化があったところだけを改変することも多く、必要な情報が埋もれてしまう、また、気づかずに誤った所見が続くことがある。例えば胸痛が1ヶ月以上続く、重篤な化学療法有害事象が3週間以上続いているが、何も対処がなされていない(つまりそれは記載が間違っている)、ということもあり信憑性が疑わしいこともある。関連して、紙カルテにおける記載は記載医が少なくとも意識していることの表れであると考えられるが、電子カルテの複写は、例えば病理所見、画像所見が範囲選択されて複写されるため意識に上っているかも不明などがある。

また、書いてあることと行っていることのギャップが予想されることから診療録の情報を絶対視すると、測定しているのが診療の質なのか、記録の質なのかというジレンマが生じてしまう。記録の質は診療の質の一側面であることは間違いないが、どこまで要求されるのかは定かではない。これまで、医師の診療録記載に関して、各種癌取り扱い規約など記載方法の基準は一部存在するが、何を書くべきで何を書かなくてもよいということは多くは慣習的なものであり基準はなかったため、必須情報について診療録記載をQIとすることは、非常に新しいことである。

d. その他

他の方法としては、標準患者を使った演技、

シナリオを使った医師の能力テスト、ビデオ撮影などの方法がある。どれも、診療録の採録よりもより手間のかかる方法であり、データベースで行うことは可能であるが、システム化することは難しい。どちらかという、教育などの目的に使用されることが多い方法である。珍しい研究であるが、Peabodyらは、標準患者、診療録採録、シナリオの使用、の3通りの診療の質評価の方法論的比較を行ったところ、診療録採録は標準患者、シナリオよりも少し低めにスコアとしては表れるということが判明している²²。これが、記録のためなのか他の原因なのかは検討を続けなければならない。

8. QIの使われ方

a. フィードバック

—通常のフィードバック。

各QIの遵守率を計算し、他の施設の遵守率ともにフィードバックする活動は海外において頻繁に行われている。アメリカ外科専門医会(American College of Surgeons)のがん委員会(Commission on Cancer)では、その認定施設に科せられているがん登録、National Cancer Data Baseを使用して、数個のQIについて、病院ごとの遵守率を計算し、全国や州内の同様の施設の中で自施設の位置を知ることができるフィードバックプログラム(Cancer Program Practice Profile Report, CP3R)を運営している^{23,24}。ここでは、各施設が与えられたIDとパスワードを使ってインターネット上でログインし、提出したがん登録のデータから、乳房温存術後の放射線療法施行、Ⅲ期結腸癌手術の術後化学療法の施行、などのQIについての遵守率を知ることができる。ここでは、その成績は、施設ごとのフィードバックであり、自施設の成績は知ることができるのに対して、並べて表示さ

れている他の施設の成績は、施設名などが伏せられておりどの成績がどの施設の者かわからないようになってきている。もちろん、成績の公表も行われていない。このプログラムにより診療の質に関する意識とともに、データの質に関する意識が高まり、がん登録の精度が向上するという効果も生まれているという（図10）。

また別の活動として、アメリカ臨床腫瘍学会（American Society for Clinical Oncology）が Quality Oncology Practice Initiative という医師ごとに参加するプログラムを運営している^{25,26}。ここでは、医師やそのグループの雇用するスタッフが、年2回、がん診療のあらかじめ定められた50程度のQIについて採録して患者ごとの適合/不適合を評価してデータをこのプログラムのサーバーに送るというものである。そこでやはり成績が計算され散会しそれぞれに対してフィードバックが行われている。QIは基礎的なものばかりであるため、常に改善を促すために持続的に改訂、追加が行われている。これは医師の自由参加であるが、専門医更新のための生涯教育の単位の要件が「質の改善活動を行っている」という項目がありその単位をこれで満たすことができることから、参加を促す動機付けにもなっている。

—達成可能なラインを含めたフィードバック

単純なフィードバックだけではなく、その方法を工夫して上位10%程度の成績を特にフィードバックして「達成可能な目標」としてフィードバックすることが有効ではないかという研究もある²⁷。Kiefらは、70の参加医師の診療を5つのQIを使って診療の質を測定し、そのフィードバックの方法を、通常行われるような、個別成績のみのフィードバ

ックする群と、上位10%の医師の平均スコアからなる「達成可能フィードバック」を合わせてフィードバックする群にランダム割り付けして、その改善具合を比較したところ、全てのQIにおいて、全体のスコアは改善をしたものの、「達成可能な目標」をフィードバックされた施設の方が全体に、改善の度合いが大きかったということを示した。これが系統的に社会的に応用されたという話はあまり耳にしないが、ちょっとした工夫として興味深いものである。

—QIにそって、診療を誘導する

QIの診療はある程度定型的であることから、それを診療に組み込んでしまうという活動、研究もある。電子化された診療システム（オーダーリングシステムや電子カルテ）では、ある特定の臨床状況、特に定期的なフォローアップ（糖尿病におけるHbA1cの測定や眼底検査など）を例えば対象患者のオーダー画面を開くと同時に表示する（リマインダ）というシステムも米国では普及している。特にこのような電子カルテシステムで有名な米国の退役軍人医療システム（Veterans Health Administration）では、（このシステムだけが要因ではないものの）近年診療の質の改善がめざましく、以前は一般よりも劣っているとされた診療の質が一般人口よりも成績がよくなったことが知られている²⁸。

研究ベースであるが、質に問題があるとされた、転倒に関する医療、尿失禁、記憶障害の3分野において

- ・スクリーニングツールの提供、
- ・医師教育セッション
- ・必要事項が実行されるように誘導するように構造化された診療録

による介入を行って、診療の質の向上を検討した研究もある²⁹。これは、電子カルテのシ

システムなどとは対照的に高度な Information Technology(IT)技術は使わずに、ほとんど人力と紙ベースのツールを使って改善を試みたということに大きな特徴があった。全体として、QI 遵守率は25%から41%に上昇したものの、まだまだ満足のいくレベルへは到達しておらず、継続的・多角的な改善活動が望まれるとされている。

一結果の公表

わが国では、QI を使って測定した結果を公表することを、当然のように論ずる向きがあるが、これは非常に慎重に検討すべき課題である。情報公開は患者が受診する施設を選択するための資料となることが期待されているが、まず第1の疑問として、患者は果たして公開情報を使うのかを検証する必要がある。これは欧米において非常に多く研究されている。最も公表されているデータ自体が過程指標を使ったデータというよりも、死亡率、合併症の率などの結果評価のデータであることが多いのは留意する必要がある。Marshall らのレビューによると、あまり患者は期待された程情報を資料として使わない、ということである³⁰。個別の研究においては、例えば心臓外科手術の外科医別死亡率が米国の数州において算定され公表されているが、その1つであるペンシルバニア州の心臓外科手術を受けた患者474名の調査では20%しかこの公表データの存在を知らず、2%しか、受診先の選択にこのデータが影響したと答えなかった。また、自分が実際の手術を受けた外科医の手術死亡率が、平均以上なのか、以下なのか、変わらないのかを正しく言えたのは1%にすぎなかったという³¹。

しかし、逆に医療者側の反応は著しい。米ウイスコンシン州の115の病院で、評価結果公表群、非公表成績測定フィードバック群、

非フィードバック群の3群(公表はある特定の保険団体と契約している病院、非公表フィードバック、非フィードバック群の2群はランダム割り付け)の群を、比較したところ、診療の質改善プログラムの数は、産科において非公表フィードバック群、非フィードバック群ではそれほど変わらないにもかかわらず、公表群においては倍以上の数のプログラムが設定されていた。循環器科においては、これほど大きな差はないものの、公表群において改善プログラムの数は最も多く、非公表フィードバック、非フィードバック群の順番で少なくなっていた。

2005年から米国の政府管掌高齢者保険(メディケア)が契約病院を対象とした「Hospital Compare」というプログラムが始まっている。ここではメディケアの契約病院から提出されたデータを使って、一定のQI(例えば、手術ではガイドラインに沿った予防抗生剤の投与が行われているか、予防抗生剤は術後24時間以内に中止されているかなど)といったQIについて成績を計算しインターネット上で公開されている。ここでは、各州、各群、ある郵便番号からの距離などの地理条件や病院の名前で病院が検索されそのそれぞれのQIについての成績が比較できるようになっている。Laschoberら³²がこのHospital Compareについて、病院の診療の質管理の担当責任者のサーベイを行ったところ、同様の結果が出ている。回答した担当者のうち23%がこの成績について外部から連絡(問い合わせや意見)を受けたことがあると回答したが、逆に病院の内部では95%以上が、内部での検討材料としてこのHospital Compareのデータが使われていると回答した。このような成績公表の活動自体が、質に関する意識を高める推進力になっているのは間違いなさそうである。

しかし、意識が高まることが即改善か、という慎重に解釈しなければならない。特に診療を評価するという活動がまだ黎明期にあるわが国では、何よりも質の改善を進めていく上での現場の臨床家の協力が必要不可欠である。成績の公開というのは臨床家にとっては非常に心理的抵抗が強いことはある意味当然であり、また、本当に改善が必要な施設や場所において抵抗が強いであろう。臨床家の協力が必要なところで、抵抗感が先行して、改善活動が悪い影響を受けるという事態は何としても避けなければならない。この辺りは既に医療の評価を行うといった社会通念が浸透している米国とは異なることを認識して考えなければならない。

さらに、成績を公表しても主にそれが影響するのは、医療者の側であり、患者がそれを使って受診先を選択するということは稀であることも注視する必要がある。つまり、成績を公表することで競争原理が働いて質の低い施設が淘汰される、ということは到底望めないのである。また、真の質向上よりも、数字あわせの成績向上を行うという動機付けが、成績の公表には強く働くということも留意する必要があるだろう。その中で最悪なのは、成績が低く出るような扱いの難しい患者（服薬指導や来院指導が守れない、多くの併存症を抱えて全てを管理するのが困難）といった患者を担当しない動機付けが働いて、施設間でこのような患者のたらい回しが起こる危険性もありえるわけである。ここでは本来地域の中の施設というのはお互いの役割を認識して協力しなければならないはずなのに、成績公表により、その関係が競争関係にすり替わり、協力が難しくなってしまうということである。

しかし、患者の中でそのような成績のデータを知りたいと考える人々は少数であって

もその要望も強いことは理解できる。また外部的な推進力がなかなか改善を実行できない病院には必要であることも理解できる。筆者の私案であるが、上記の問題点をできるだけ避けつつ公表を進めるには、施設ごとの成績公表などではなく、本来協力し合う単位、例えば地域ごとの公表がその中間点として妥当ではないかと考える。医療のサービスは基本的に地域が主体であることから、各施設が地域の医療を担う一員として改善を施行するにはこのような工夫が必要ではなかるうか。

9. まとめと結語

以上、診療ガイドラインと関連の深い、過程評価を中心とした Quality Indicator の考え方、作成方法、その海外における使用例について、概観した。わが国での動きはまだまだこれからである。しかし、診療ガイドライン作成と同様、QI の内容は非常に専門性の高いものであり、専門家が主導して自律的に考えていくことが必要であるのはいうまでもない。一方で、診療・医療を評価すべきである、という社会の要請は非常に強いものである。その要請に専門家が応える姿勢を示せるのか、また応えていけるのかで、医療崩壊という流行語に代表される、社会と医療の隔絶から医療が再生していけるのかが決まるだろう。

このような評価活動を行う上では、あくまで評価が改善を目的としていることを強く意識することは大変重要である。ともすると、ランキング、患者が医療機関を選ぶための情報を提供するということが目先の目標になってしまうリスクは常にある。しかし、本来、患者が医療機関を選ばなければならないというのは不幸な状況である。医師・患者の相性や、個々のある程度能力の違いがあるのは

自然なことなのかもしれないが、どこへ行っても一定レベルの医療が保障されている、という状態まで、改善を重ねていく、つまり患者が医療機関を「選ばなくてもよい」状況を作ることを考えなければならない。それは医療者が、社会からその責任を、「免許」という独占の特権を付与されることで、同時に期待されている最重要課題ではないだろうか。

参考文献

1. Donabedian A. The Definition of Quality and Approaches to its Assessment. In. Ann Arbor: Health Administration Press; 1980.
2. The Institute of Medicine. Medicare: A Strategy for Quality Assurance. Washington D.C: National Academy Press; 1990.
3. 多尾清子. 統計学者としてのナイチンゲール. 東京: 医学書院; 1991.
4. Donabedian A. The end results of health care: Ernest Codman's contribution to quality assessment and beyond. *Milbank Q* 1989; 67: 233-56; discussion 57-67.
5. 安藤優一郎. 江戸の養生所. 東京: PHP 研究所; 2005.
6. Iezzoni LI. The Risks of Risk Adjustment. *JAMA* 1997;278:1600-7.
7. Iezzoni LI. Risk Adjustment. Chicago, IL: Health Administration Press; 2003.
8. Keeler EB, Rubenstein LV, Kahn KL, Draper D, Harrison ER, McGinty MJ, Rogers WH, Brook RH. Hospital characteristics and quality of care. *JAMA* 1992;268:1709-14.
9. Hanlon JT, Schmader KE, Samsa GP, Weinberger M, Uttech KM, Lewis IK, Cohen HJ, Feussner JR. A method for assessing drug therapy appropriateness. *J Clin Epidemiol* 1992; 45:1045-51.
10. Caplan RA, Posner KL, Cheney FW. Effect of outcome on physician judgments of appropriateness of care. *JAMA* 1991; 265: 1957-60.
11. McGlynn EA, Asch SM, Adams J, Keesey J, Hicks J, DeCristofaro A, Kerr EA. The quality of health care delivered to adults in the United States. *N Engl J Med* 2003; 348:2635-45.
12. Wenger NS, Roth CP, Shekelle P. Introduction to the assessing care of vulnerable elders-3 quality indicator measurement set. *J Am Geriatr Soc* 2007; 55 Suppl 2:S247-52.
13. The Institute of Medicine. Crossing the Quality Chasm. Washington D.C: National Academy Press; 2001.
14. Wennberg J, Gittelsohn. Small area variations in health care delivery. *Science* 1973; 182:1102-8.
15. AGREE Research Trust. Appraisal of Guidelines Research and Evaluation Instrument. 2008. (Accessed Feb 13, 2009, at <http://www.agreetrust.org/>.)
16. Kerr EA, Asch SM, Hamilton EG, McGlynn EA. Quality of Care for General Medical Conditions. Santa Monica: RAND Health; 2000.
17. Fitch K, Bernstein SJ, Aguilar MD, Burnand B, LaCalle JR, Lazaro P, van het Loo M, McDonnell J, Vader JP, Kahn KL. The RAND/UCLA Appropriateness Method User's Manual: RAND; 2001.
18. Wenger NS, Solomon DH, Roth CP, MacLean CH, Saliba D, Kamberg CJ, Rubenstein LZ, Young RT, Sloss EM, Louie R, Adams J, Chang JT, Venus PJ, Schnelle JF, Shekelle PG. The quality of medical care provided to vulnerable community-dwelling older patients. *Ann Intern Med* 2003; 139: 740-7.
19. Nolan T, Berwick DM. All-or-none

- measurement raises the bar on performance. *JAMA* 2006; 295:1168-70.
20. Doran T, Fullwood C, Gravelle H, Reeves D, Kontopantelis E, Hiroeh U, Roland M. Pay-for-performance programs in family practices in the United Kingdom. *N Engl J Med* 2006; 355:375-84.
 21. Desch CE, McNiff KK, Schneider EC, Schrag D, McClure J, Lepisto E, Donaldson MS, Kahn KL, Weeks JC, Ko CY, Stewart AK, Edge SB. American Society of Clinical Oncology/National Comprehensive Cancer Network Quality Measures. *J Clin Oncol* 2008; 26:3631-7.
 22. Peabody JW, Luck J, Glassman P, Dresselhaus TR, Lee M. Comparison of vignettes, standardized patients, and chart abstraction: a prospective validation study of 3 methods for measuring quality. *JAMA* 2000; 283:1715-22.
 23. Bilimoria KY, Stewart AK, Winchester DP, Ko CY. The National Cancer Data Base: a powerful initiative to improve cancer care in the United States. *Ann Surg Oncol* 2008; 15: 683-90.
 24. 東尚弘. 米がん登録を利用した診療の質向上活動. *癌と化学療法* 2008; 35: 1445-9.
 25. Jacobson JO, Neuss MN, McNiff KK, Kadlubek P, Thacker LR, 2nd, Song F, Eisenberg PD, Simone JV. Improvement in oncology practice performance through voluntary participation in the Quality Oncology Practice Initiative. *J Clin Oncol* 2008; 26:1893-8.
 26. McNiff KK, Neuss MN, Jacobson JO, Eisenberg PD, Kadlubek P, Simone JV. Measuring supportive care in medical oncology practice: lessons learned from the quality oncology practice initiative. *J Clin Oncol* 2008; 26:3832-7.
 27. Kiefe CI, Allison JJ, Williams OD, Person SD, Weaver MT, Weissman NW. Improving quality improvement using achievable benchmarks for physician feedback: a randomized controlled trial. *JAMA* 2001; 285: 2871-9.
 28. Asch SM, McGlynn EA, Hogan MM, Hayward RA, Shekelle P, Rubenstein L, Keeseey J, Adams J, Kerr EA. Comparison of quality of care for patients in the Veterans Health Administration and patients in a national sample. *Ann Intern Med* 2004; 141: 938-45.
 29. Wenger NS, Roth CP, Shekelle PG, Young RT, Solomon DH, Kamberg CJ, Chang JT, Louie R, Higashi T, Maclean CH, Adams J, Min LC, Ransohoff K, Hoffing M, Reuben DB. A Practice-Based Intervention to Improve Primary Care for Falls, Urinary Incontinence, and Dementia. *J Am Geriatr Soc* 2009.
 30. Marshall MN, Shekelle PG, Leatherman S, Brook RH. The public release of performance data: what do we expect to gain? A review of the evidence. *JAMA* 2000; 283: 1866-74.
 31. Schneider EC, Epstein AM. Use of public performance reports: a survey of patients undergoing cardiac surgery. *JAMA* 1998; 279: 1638-42.
 32. Laschober M, Maxfield M, Felt-Lisk S, Miranda DJ. Hospital response to public reporting of quality indicators. *Health Care Financ Rev* 2007; 28:61-76.

図1: 診療の質評価の3つの視点

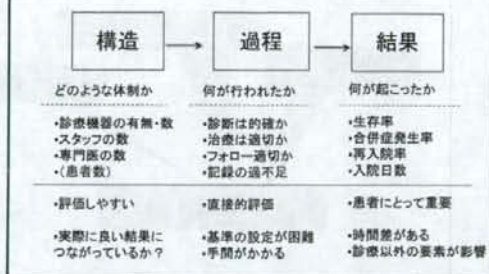


図2 結果評価のイメージ

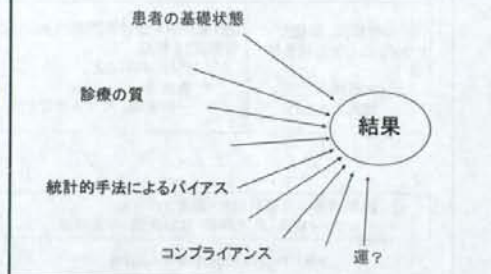


図3 過程評価のイメージ



図4 評価対象となる診療のタイプ

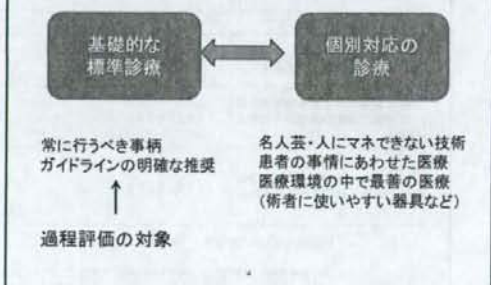


図5 用語の違い

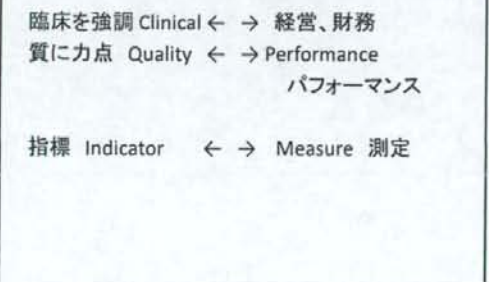


図6 ガイドライン推奨と診療指標の関係

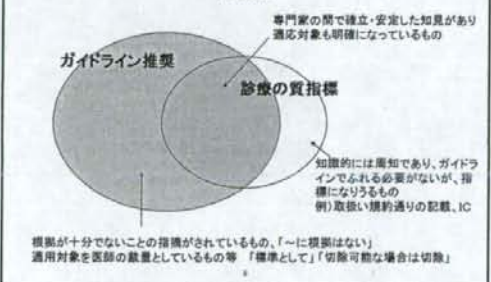


図7 Quality Indicatorの作成方法

エビデンスに基づく客観的コンセンサス検討

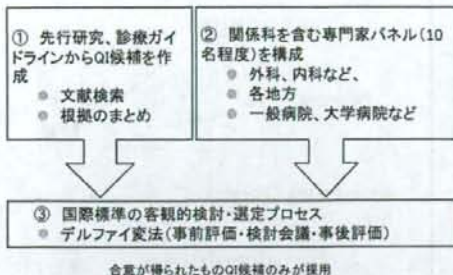


図8 QI候補の情報源

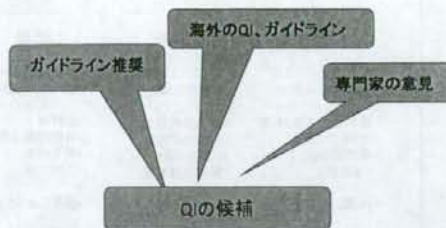


図9 QI候補評価シートの例

QI候補	QIとしての適切性	コメント
手術を受ける患者には、手術の利益と不利益の両方が説明される	1 2 3 4 5 6 7 8 9	
がんと診断された患者は病歴に関して評価がなされ診療録に記載される	1 2 3 4 5 6 7 8 9	
手術以外の治療が行われる患者は治療開始前に確定診断がなされる	1 2 3 4 5 6 7 8 9	

1=QIとしてきわめて不適切、 9=QIとしてきわめて適切

(注) QI候補の内容は例であり、実際のものとは異なります。

図10 診療の質⇔データの質

- 診療の質の問題点が明らかになる
- ↓
- 診療の質改善への関心が高まる
- ↓
- 関係者がデータを真剣に検討する
- ↓
- 病院でのがん登録への協力体制の整備される
- ↓
- より信頼性のある情報が収集される

ステロイド長期処方患者における骨粗鬆症の管理と治療の現状 : 診療ガイドラインの遵守状況とその関連因子

研究協力者 桐ヶ谷大淳 京都大学大学院医学研究科社会健康医学系専攻
健康情報学分野

研究代表者 中山健夫 同上

趣旨:

骨粗鬆症の発症原因の一つとして副腎皮質ステロイド薬(ステロイド)がある。ステロイド性骨粗鬆症への早期対応の重要性が認識されるようになり、日本でも「ステロイド性骨粗鬆症の管理と治療のガイドライン」が2004年に策定された。しかし、臨床現場におけるガイドラインの遵守状況について大規模には分かっていない。本研究ではステロイド性骨粗鬆症の管理と治療のガイドラインの遵守状況を調査し、年齢や性別、ステロイド処方量、診療科、施設規模といった要因との関連性について検討を行った。

健康保険組合のレセプト(診療報酬請求明細書)データベースを用いてガイドラインの対象基準とみなされる症例を抽出した(2004年1月~2007年12月)。ステロイド処方期間中の骨粗鬆症治療薬の処方と骨密度測定の有無からガイドライン遵守の有無を判断した。年齢や性別、ステロイド処方量、診療科、施設規模などの要因とガイドライン遵守の有無との関連を解析した。ガイドライン遵守の判定基準が違うプレドニゾン(PSL)5mg/日以上と5mg/日未満に分けて層別解析を行った。

基準を満たす対象2368人のガイドライン遵守割合は全体で23.3%であった。PSL換算5mg/日未満の症例では8.3%であり、経過観察のための骨密度測定の実施頻度は非常に少なかった。全例で骨粗鬆症の薬物療法が推奨されるPSL換算5mg/日以上症例では30.5%であり、海外の報告よりも低い割合であった。若年者、男性、ステロイド処方量がより少ない場合にガイドライン遵守割合が低かった。診療体系では内科系に比べて外科系、耳鼻咽喉科で遵守割合が低く、小規模施設ほど遵守割合が低かった。PSL5mg/日未満の症例では女性、整形外科、小規模病院のオッズ比が大きく、耳鼻咽喉科と皮膚科のオッズ比が小さかった。PSL5mg/日以上症例では耳鼻咽喉科で有意差がなくなった以外は、全体の解析結果と同様の傾向がみられた。

以上から、ステロイド長期処方患者における骨粗鬆症の管理と治療は、ガイドラインの推奨と照らして十分とは言えない。ガイドライン順守割合が低い患者層、診療体系を中心に啓発を行うことで、診療の質を改善することが期待できる。

今後、他の診療ガイドラインについても、レセプトデータベースを用いた「ガイドライン・診療ギャップ」の実状と関連要因解明の可能性を追求するものである。

1. 研究の背景

骨粗鬆症は、骨強度低下を特徴とし骨折のリスクが増大しやすくなる骨格疾患と定義されている¹⁾。日本の高齢者における要介護原因

のうち転倒骨折が第3位(10.8%)を占め(平成16年国民生活基礎調査)、将来の健康のためにも骨粗鬆症や骨折の予防は重要な問題である。閉経後や高齢者にみられる原発性

骨粗鬆症とは別に、二次的に骨粗鬆症を引き起こす代表的なものとして副腎皮質ステロイド薬(ステロイド)がある。ステロイドは強力な抗炎症や免疫抑制などを目的によく使用されている薬剤だが副作用も多く、骨粗鬆症は重大なものの一つである。

1996年の米国リウマチ学会の調査で、全米の骨粗鬆症患者の20%がステロイドによるものであり、ステロイド長期投与患者の25%が骨折するという状況が明らかになった²⁾。経口ステロイド使用開始後3~6ヵ月で新規椎体骨折の発生率は最大になるという報告があり³⁾、現在では早期からの対応の重要性が認められている。1994年頃から骨粗鬆症治療薬の大規模臨床試験が次々と行われ、ステロイド性骨粗鬆症においても最も効果的なものとしてビスフォスフォネート製剤⁴⁾、それより劣るが活性型ビタミンD₃やビタミンK₂製剤でも骨折予防効果が示されている^{5) 6)}。欧米では1996年よりステロイド性骨粗鬆症の管理ガイドラインが発表され、その後改訂も行われてきた。日本でも2004年に「ステロイド性骨粗鬆症の管理と治療のガイドライン」が初めて策定された⁷⁾。18歳以上で経口ステロイドを3ヵ月以上使用中または使用予定の患者を対象とし、ごく初期からの管理と治療を重視したものになっている(図1)。

ガイドラインが発表されて5年近く経過したが、臨床現場におけるガイドラインの遵守状況について単一施設の報告が散見されるのみで^{8) 9)}、大規模には調査されていない。ステロイドは診療科を問わず多くの分野で使用されている薬剤のため、他のガイドラインよりも利用に関する実態を把握しにくい。外科系でステロイド性骨粗鬆症への認識が低いという報告があり¹⁰⁾、診療科によってガイドラインの遵守状況が異なる可能性がある。また骨粗鬆症自体は閉経後女性や高齢者に多い疾患で

あり、男性や若年層のステロイド長期処方患者では管理や治療が十分に行われていない可能性もある。

エビデンスと実際の診療にはギャップがあることが近年明らかになっている¹¹⁾。診療の質を向上させるにはエビデンスをつくることだけではなく、それを伝えることや使うことも重要である。診療ガイドラインによる標準的治療の普及は、エビデンスを伝える手段の一つである。ガイドラインは診療の支援を目的としており無批判に遵守されるべきものではないが、あまりに低い遵守割合は診療の質に問題のあることが推測される。ガイドラインの遵守状況とその関連因子を探ることは、ガイドラインと診療とのギャップがどこにあるか知る手がかりとなる。健康保険組合のレセプトデータベースを用いて、それを調査することにした。

2. 目的

ステロイド性骨粗鬆症の管理と治療のガイドラインの遵守状況を調査し、年齢や性別、ステロイド処方量、診療科、施設規模といった要因と関連があるか明らかにする。

3. 方法

1) 研究デザイン

レセプト(診療報酬請求明細書)データを用いた横断研究

2) 使用したデータベース

株式会社日本医療データセンター(Japan Medical Data Center Co. Ltd)(JMDC)が保有する7健康保険組合(累積被保険者数53万件)のレセプトデータベース(JMDC-MDB)を用いた。JMDC-MDBはハッシュ値などにより匿名化されているが、個を識別可能な暗号化が行われている。独自の匿名化名寄せプログラムおよび傷病名辞書によりデータの質を担保している¹²⁾。

3) 研究対象期間

2004年1月～2007年12月診療分

4) 研究対象者の選択基準、除外基準

ガイドラインの対象は、経口ステロイドを3ヵ月以上使用中または使用予定の症例である。しかしレセプトデータから使用予定なのかは分からないため、継続して3ヵ月以上使用中の症例を選択基準とした。まずJMDC-MDBから、研究対象期間中に経口ステロイドを28日以上処方されている月が全体で3ヵ月以上ある18歳以上の症例を抽出した。その中で、継続処方が90日未満の症例は除外した。次に、体内の副腎皮質ホルモンが不足し補充が必要である下垂体機能低下症や副腎機能不全症の病名のある症例、緩和ケアなど特殊な状況が想定される悪性腫瘍が疑わしい症例、ステロイド処方量が不明な症例は除外した。

5) 背景情報

研究対象者の背景情報を得るためJMDC-MDBから生年と性別、経口ステロイド処方内容、診療科、施設規模を抽出した。生年およびデータ上で経口ステロイド処方が開始された年月から、年齢を算出した。経口ステロイド薬の種類からプレドニゾン(PSL)換算量に統一し、ステロイド処方量を算出した。処方量に増減がある場合、処方期間中の連続90日の平均値が最大となる数値を採用した。複数の診療科で経口ステロイドが処方されている場合、主となって処方している診療科を採用した。施設規模は診療所(0-19床)、小規模病院(20-199床)、中規模病院(200-499床)、大規模病院(500床以上)の4つに区分した。

6) 主たるアウトカム指標

ガイドラインの遵守を主たるアウトカム指標とした。ガイドラインのフローチャートでは、すでに脆弱骨折があるか治療中に骨折がある、または骨密度測定で若年成人平均(YAM)

80%未満の場合には、骨粗鬆症治療として薬物療法(第1選択:ビスフォスフォネート製剤、第2選択:活性型ビタミンD₃またはビタミンK₂)を行うことが推奨されている。しかし、レセプトデータでは傷病名と実際に罹患した疾患は必ずしも同一ではなく、検査をしたことは分かっても検査結果までは分からない。そのため、本研究ではステロイド処方量と骨粗鬆症治療薬の処方、骨密度測定の有無からガイドライン遵守状況を評価することにした(図2)。

海外のメタ分析で骨折率はステロイド用量依存性に上昇し、PSL5mg/日以上が骨折リスク増加の閾値という報告がある³⁾。それに準じて日本のガイドラインでは、PSL5mg/日未満なら経過観察として骨密度測定と胸腰椎X線撮影を定期的(6ヵ月～1年毎)に行うこと、PSL5mg/日以上なら骨密度測定値にかかわらず全例で薬物療法を行うことを推奨している。そこでPSL5mg/日未満の症例では、ステロイド処方期間中に骨粗鬆症治療薬が処方されている場合、脆弱骨折歴か骨密度測定値から薬物療法が開始されたときみなしガイドライン遵守ありとした(①)。骨粗鬆症治療薬が処方されていない場合、経過観察として骨密度測定が定期的に行われていればガイドライン遵守ありとした(②)。PSL5mg/日以上の症例では、ステロイド処方期間中に骨粗鬆症治療薬が処方されていればガイドライン遵守ありとした(③)。

胸腰椎X線撮影はデータの項目から他部位の撮影と区別困難なため、判定基準は緩くなるが骨密度測定のみで経過観察がされているか評価した。薬物療法に関して、第1,2選択薬以外の骨粗鬆症治療薬のみ処方されている症例はガイドライン遵守ありに含めなかった。

4. 解析

はじめに研究対象者の人数を把握し、背景として患者属性(性別、年齢、ステロイド処方量)と医療機関属性(診療科、施設規模)を記述した。ステロイド処方量と骨粗鬆症治療薬の処方、骨密度測定からガイドライン遵守の有無を決定した。年齢や性別、ステロイド処方量、診療科、施設規模といった要因と、主なアウトカム指標であるガイドライン遵守の有無との関連について、まず単変量解析を行った。次にガイドライン遵守の有無を目的変数、それぞれの要因を説明変数として、ロジスティック回帰モデルによる多変量解析を行った。さらに、ガイドライン遵守の判定基準が違う PSL5mg/日以上と5mg/日未満に分けて層別解析を行った。両側検定、有意水準は0.05、数量的データの解析には Stata Ver.10 を使用した。

5. 倫理的配慮

本研究の対象となるデータは連結不可能匿名化された既存資料である。従って、解析データセットから個人を特定し得る情報は同定されない。得られた情報は本研究の目的以外には用いない。

本研究で用いるデータは JMDC より無償で提供を受け、保健医療情報分析活用研究会の活動の一部として実施した。研究会は JMDC と利害関係を有していない。

6. 結果

1) 研究対象者の抽出

JMDC-MDB から、研究対象期間中に18歳以上で経口ステロイドを28日以上処方されている月が全体で3ヵ月以上ある症例を条件として検索したところ、3058人が抽出された。そのうち継続処方が90日未満である449人は除外した。次に、傷病名と医薬品データから下垂体機能低下症54人、副腎機能不全症37人、悪性腫瘍の疑いが強い147人、ステロイド処方

量が不明な3人は除外した。残った2368人を本研究における対象者とした(図3)。

2) 研究対象者の背景

研究対象者2368人の背景を表1に示す。年齢は平均45.9歳(最高98歳、最低19歳)、女性が53.5%と半数以上であった。ステロイド処方量の平均値はPSL6.6mg/日、約1/3は5mg/日未満であった。診療科は一般内科にリウマチ科など内科系専門科をあわせて内科系、一般外科に脳神経外科など外科系専門科をあわせて外科系とした。症例数が多い整形外科、皮膚科、耳鼻咽喉科は単独で評価した。内科系が57.4%と半数以上を占めた。施設規模では診療所(0-19床)が半数を占めた。ステロイド処方期間中に第1,2選択の骨粗鬆症治療薬が処方されているのは22.9%であった。一人に複数薬の処方例もあるが、活性型ビタミンD₃の処方が最も多く、ビタミンK₂はあまり処方されていないかった。骨密度測定は骨粗鬆症未治療者には4.3%しか行われておらず、経過観察としての骨密度測定はほとんど行われていないことが明らかになった。

3) ガイドラインの遵守状況

PSL5mg/日未満の症例のうち、ステロイド処方期間中に骨粗鬆症治療薬が処方されている56人(7.3%)、骨粗鬆症治療薬は処方されていないが経過観察として定期的(6ヵ月~1年毎)に骨密度測定が行われている8人(1.0%)、あわせて64人(8.3%)をガイドライン遵守ありとした。PSL5mg/日以上の症例のうち、ステロイド処方期間中に骨粗鬆症治療薬が処方されている487人(30.5%)をガイドライン遵守ありとした。全体ではあわせて551人(23.3%)をガイドライン遵守ありとした(表2)。

4) 要因ごとのガイドライン遵守割合とオッズ比(単変量解析および多変量解析)

要因ごとのガイドライン遵守割合、単変量解析およびロジスティック回帰モデルによる多変

量解析で算出したオッズ比を表3に示す。年齢は高いほど、性別では女性で調整後もオッズ比が大きかった。ステロイド処方量は PSL5mg ずつのカテゴリーに分けて評価したが、調整すると小さくなるものの用量依存性にオッズ比は大きくなった。診療体系では診療科による違いもみられ、内科系より外科系と耳鼻咽喉科で調整後もオッズ比は小さかった。皮膚科でも調整前には有意差があったが、調整後には無くなった。施設規模では調整すると小さくなるものの、規模が大きいほどオッズ比は大きくなった。

5) ステロイド処方量による層別解析

ガイドライン遵守の判定基準が違う PSL5mg/日以上と 5mg/日未満に分けて層別解析を行い、それぞれの調整オッズ比を算出した(表4)。PSL5mg/日未満では女性、整形外科、小規模病院のオッズ比が大きくなり、耳鼻咽喉科と皮膚科のオッズ比が小さくなる傾向がみられた。PSL5mg/日以上では耳鼻咽喉科で有意差がなくなった以外は、全体の解析結果と同じような傾向がみられた。

7. 考察

1) 本研究の結果について

米国での診療の質に関する研究では、急性期や慢性期、予防医療ともに推奨されている診療がされているのは 50%強であった¹¹⁾。これと比較すると、ガイドライン遵守割合が全体で 23.3%というは決して良い数字ではない。PSL5mg/日未満の症例では特に割合が低く、経過観察のための定期的な骨密度測定はほとんどされていないことが分かった。ステロイド処方量が少なくても長期処方する場合は、定期的な骨密度測定など経過をみるように啓発するべきである。全例で骨粗鬆症の薬物療法が推奨される PSL5mg/日以上での遵守割合は 30.5%であった。1996年にガ

イドラインが策定された米国では、PSL5mg/日以上処方されている患者での骨粗鬆症治療薬の処方割合は 1998年からの2年間は約9%、2003年からの3年間は約47%であった^{13) 14)}。異なるデータベースであり背景に骨粗鬆症治療薬自体の処方の伸びもあるが、数年でステロイド性骨粗鬆症の治療状況が改善している。米国と比較すると日本での薬物療法の割合はまだ低いが、ガイドラインが適切に普及していけば増加していくことが期待できる。

要因との関連について、遵守割合の高い患者層はより高齢、女性、ステロイド処方量のより多い症例であった。これは原発性骨粗鬆症や脆弱骨折歴の多い患者層であり、アウトカムが過大評価されている可能性もあるが、骨粗鬆症の検査や治療を抵抗なく行いやすいことを反映しているのかもしれない。診療科では内科系よりも外科系と耳鼻咽喉科で遵守割合が低かった。骨粗鬆症患者を診ることが少ない診療科でもあり、ガイドラインの啓発により力を入れていくべきである。また施設規模では規模が小さいほど遵守割合が低く、ガイドラインが十分に浸透していないだけではなく、他診療科の併診や検査をすぐにできないといった弊害も考えられる。層別解析の PSL5mg/日群における要因ごとの結果は、ステロイド処方量が少ないため、ステロイド処方の有無に関係ない全体における骨粗鬆症治療の傾向に近いのではないかと推測される。

2) レセプトデータの特徴、本研究への応用

レセプトは患者の基本情報や傷病名、診療行為、薬剤名、診療報酬請求額などが原則1医療施設ごと、1ヵ月単位で記載されている。それを集積したレセプトデータベースは、多くの調査研究で使われ始めている¹⁵⁾。特長として、自動的に大きなサンプルサイズのデータが集積されるため経済的であり、国や地域間

の比較研究などマクロな視点での分析も可能である。毎月集積されるため、時間経過の追跡や長期間における傾向の変化なども調査できる。研究参加の意思に関係なく現状を調査できるため、報告バイアスが入らない。また、JMDC-MDB では患者の受診したすべての医療機関における検査や処方内容などを把握できる。複数の施設や診療科に受診してもデータが抜け落ちることがないため、評価対象の医療機関で行われている診療しか調査していないという落とし穴に陥ることはない。一方、研究指向型データではないため制約される点もある。検査や処方のために便宜的な傷病名が記載される場合があり、傷病名と実際に罹患した疾患は必ずしも同一ではない。検査をしたことは分かっても結果までは分からない。解析可能な状態にするためのデータセットの再構築やデータクリーニングに時間がかかり、場合によってはデータにあうように研究目的を再構築する獨創性が必要なこともある¹⁶⁾。長短あるものの国民皆保険の日本ではほとんどの診療情報はレセプトに集約されるため、その限界を正しく理解して利用すればレセプトデータベースは価値の高い資料となる。

本研究ではレセプトデータベースを用いることで、大規模なサンプルサイズを得ることができ、単一施設の調査では困難な診療体系ごとの分析も行うことができた。バイアスを減らすために、対象者を傷病名からではなく、信頼性の高い項目であるステロイド処方量と期間から選択した。主なアウトカム指標であるガイドライン遵守も、信頼性の高い項目で判断できるように対処した。日本ではここ10年ほどの間にEBMの手法を用いた診療ガイドラインが次々に作成され、Minds (<http://minds.jcqh.or.jp/index.aspx>)のようにそれを普及する事業も行われてきた。しかし、診療ガイドラインと実際の診療とのギャップがど

れくらいあるかについての報告はまだほとんどなく、各領域で調査されているところである。レセプトデータを使うことで、他のガイドラインにおいても実際の診療の現状や今後長期間における変化を調査できるかもしれない。

3) 本研究の限界と今後の課題

JMDC-MDB は健康保険組合のみのデータであり、日本の一般的な患者層より若年者が多い。抽出バイアスとなる可能性があるが、若年者でもステロイド性骨粗鬆症の早期管理、治療の重要性に変わりはない。国民健康保険のデータの利用は現段階では難しい状況であるが、2011年にはレセプトのオンライン請求義務化も控えており今後の整備が望まれる。ガイドラインの対象者のうち3ヵ月以上使用予定のものに対する状況は分からないが、処方期間が短くなるため遵守割合は下がると予測される。ガイドライン策定後の現状をみるのが目的であったが、ステロイド新規処方例が少ないため、継続処方例を含めて横断的に解析を行った。継続処方例ではステロイド開始前から骨粗鬆症の薬物療法がされている場合もあり、アウトカムが過大評価されている可能性がある。また、既存資料のみの研究であるためガイドラインに遵守していない場合、それがどういう理由によるかは分からない。遵守ありとしても、本当にガイドラインを知っていて遵守したかは分からない。診療の現状を大規模に知ることにはできたが、ガイドライン遵守の有無がどういう理由によるかを同定するには、個別の事例にあたるなど更なる研究が必要である。

8. 結論

ステロイド長期処方患者における骨粗鬆症の管理と治療の現状は、まだ不十分なものであった。ガイドライン遵守割合が低い要因として若年者や男性、ステロイド処方量がより少ない症例、外科系、耳鼻咽喉科、規模のより

小さい施設が考えられた。ガイドラインとのギャップが大きい患者層や診療体系を中心に啓発を行うことで、診療の質を改善することが期待できる。

9. 謝辞

データベースを提供して頂いた日本医療データセンター、保健医療情報分析活用研究会の皆様、御指導いただいた中山健夫先生ならびに健康情報学教室、指導教官の皆様、MCR 同期の皆さん、MCR コースで学ぶ機会を下された地域医療振興協会の皆様に心より感謝いたします。

10. 参考文献

- 1) Osteoporosis Prevention, Diagnosis, and Therapy. NIH consensus Statement 2000; 17:1-45.
- 2) American college of rheumatology task force on osteoporosis guidelines. Recommendations for the prevention and treatment of glucocorticoid-induced osteoporosis. Arthritis Rheum 1996; 39: 1791-801.
- 3) van Stata TP, Leufkens HG, Cooper C, et al. The epidemiology of corticosteroid induced osteoporosis: a meta-analysis. Osteoporos Int 2002;13:777-87.
- 4) Homik J, Cranney A, Shea B, et al. Bisphosphonates for steroid induced osteoporosis. Cochrane Database Syst Rev 2000:CD001347.
- 5) de Nijs RN, Jacobs JW, Algra A, et al. Prevention and treatment of glucocorticoid-induced osteoporosis with active vitamin D₃ analogues: a review with meta-analysis of randomized controlled trials including organ transplantation studies. Osteoporosis Int 2004; 15: 589-602.
- 6) 田中郁子, 大島久二. ステロイド性骨粗鬆症の診断と治療に関する縦断研究. 診断・治療指針への予備的検討. Osteoporos Jpn 2003;11:11-4.
- 7) Nawata H, Soen S, Takayanagi R, et al. Guidelines on the management and treatment of glucocorticoid-induced osteoporosis of the Japanese Society for Bone and Mineral Research (2004). J Bone Miner Metab 2005;23:105-9.
- 8) 宮里照巳. ステロイド治療におけるビスフォスフォネート併用調査. Osteoporosis Jpn 2007;15:740.
- 9) 黒木昌幸, 梅北邦彦, 楠元規生, 他. 当科におけるステロイド性骨粗鬆症の治療実態. 日本内科学会雑誌 2008;97:149.
- 10) 紅林昌吾, 合屋佳世子, 北村哲宏, 他. ステロイド療法の合併症に関する医師の意識と管理状況. ステロイド性骨粗鬆症のリスクマネージメントの重要性. Osteoporosis Jpn 2004;12:377-83.
- 11) McGlynn EA, Asch SM, Adams J, et al. The quality of health care delivered to adults in the United States. N Engl J Med 2003;348:2635-45.
- 12) 木村真也. レセプトにおける匿名化名寄せ技術と傷病名辞書. 公衆衛生. 2007;71(11):939-42.
- 13) Ettinger B, Chidambaram P, Pressman A. Prevalence and determinants of osteoporosis drug prescription among patients with high exposure to glucocorticoid drugs. Am J Manag Care 2001;7:597-605.
- 14) Chitre MM, Hayes W. 3-year results of a member and physician intervention to reduce risk associated with glucocorticoid-induced osteoporosis in a health plan. J Manag Care Pharm 2008; 14:281-90.
- 15) 小林廉毅. 公衆衛生領域の医療経済データと個人情報保護. 学術の動向 2007.12:54-58.
- 16) Leiyu Shi. Health services research