厚生労働科学研究費補助金（臨床研究等 ICT 基盤構築研究事業）
分担研究報告書

退院サマリの品質の自動評価に関するパイロット研究

研究分担者　森田 瑞樹
（岡山大学 大学院医歯薬学総合研究科 准教授）

研究要旨

退院サマリの自動生成技術の実現を目指し，昨年度までの調査結果の整理およびそれを踏まえた退院サマリの自動評価のコンセプト検証を実施した。退院サマリを生成および評価するためには，どのような退院サマリを生成しなくてはならないかを示す「理想的な退院サマリ」の定義が必要となる。昨年度に理想的な退院サマリについて言及をした国内外の文献調査を実施し，今年度はその結果を整理した。退院サマリの記載に関して「記載すべき項目」および「定性的な要件事項」を抽出し，前者として 38 項目を得た。このうちの 14 項目はカルテの構造化データより抽出できるものの，残りの 24 項目はカルテの自由記載より抽出してサマリとして文章を作成や要約をする必要があるものと考えられた。また，後者として 9 項目を得た。このうちの 5 項目は医学的な知識・経験がないと採点が難しいものの，残りの 4 項目は形式的に判断をすることが可能と考えられた。この 4 項目のうち医学用語辞書を必要としない 3 項目について文章の特徴を用いて自動評価することを試み，人による評価との相関および点数の分布を踏まえて自動評価の可能性を考察した。退院サマリの自動評価手法の確立に向けた今後の課題が明らかとなった。

A．はじめに

近年，レセプトや DPC などの大規模な医療データ（いわゆる医療ビッグデータ）を用いた分析が研究や病院経営などのために盛んに実施されている一方で，カルテに文章として記載された情報の利活用は進んでいない。カルテの文章の活用を容易にするためには，記載がある程度は標準化されていることが望ましい。そこで本研究では，退院サマリ（退院時要約）の作成を自動化することにより記載内容を標準化することを目指している。

退院サマリとは，入院していた患者が退院する際に，入院に至った経緯から入院中の経過，および退院後の治療方針などをまとめたものであり，担当医などによって記載される。診療行為を大きく入院と外来に分けると，入院においては外来と比べて短期間に多くの医療行為が実施されるため，カルテの記載量は多くなる。退院して外来に移行する際などに，その内容を効率的に共有するためには入院記録をまとめた退院サマリが効果を発揮すると期待される。現在，医療機関の機能分化が進められており，異なる医療機関や種類の異なる医療施設（病院と介護施設など）でのスムーズな連携を行うために，今後，退院サマリの役割は増していくものと想定される。

退院サマリを自動動成する方法は自明ではない。たとえば，入院カルテからいくつかの文を抽出して組み合わせたり，退院サマリの雛形に必要な情報を入院カルテより抽出もしくは推定して埋めたりするなど，いくつかの方法が考えられる。いずれにしても，どのような退院サマリが望ま

しいのかを明らかにすることが，自動生成の指標になるものと思われる。また，生成した退院サマリを自動で評価することができれば，その評価結果に基づいてより望ましい退院サマリを選び出すことができるはずである。このため，退院サマリの自動生成のために必要な事項として，理想の退院サマリとは何かを明確に定義すること，生成した退院サマリを自動で評価できること，があると考えた。

昨年度，望ましい退院サマリに関する文献の調査を行い，結果として 51 報の文献を得た（英語：48 報，日本語：3 報）。今年度は，51 報の文献の内容を整理し，退院サマリに記載すべき項目として「記載すべき項目」および「定性的な要件事項」を抽出した。次いで，文献から抽出して整理した要件に基づいて退院サマリを評価するための方法を作成することを試みた。文章の自動評価方法は，正解となる文章との比較に基づいて評価をするものと正解を用いないで評価をするものがあるが，退院サマリの評価においては実臨床で使用することを考慮すると後者の方法しか取り得ない。本研究では，定性的な要求事項を自然言語処理を用いて簡便に評価できることを指向し，退院サマリ 50 報（人工的に作成したダミーデータ）に対して適用した。なお，定性的な要求事項には医学的な知識が必要なものとそうでないものとがあるが，本研究では後者のみを対象とした。

B．研究方法

望ましい退院サマリに関して書かれた文献 51 報から，記載すべき項目と定性的な要求事項を抽出して分類整理した。記載すべき項目は，英国 Academy of Medical Royal Colleges（AoMRC）による退院サマリの項目分類（2013 年）に基づいて退院サマリに記載すべき項目として整理した。また，文献に登場した退院サマリの記載における定性的な要求事項を分類した。

定性的な要求事項のうち医学的な知識が必要とされない項目として文や単語の特徴を評価した。評価の対象となったものは，「文法が正しい」，「言葉遣いが適切である」，「簡潔である」の 3 項目（9 項目中）であった。文章校正用のソフトウェアを使用し，そこから得られる解析結果を定性的な要求事項に合うようにスコア化した。文章校正用のソフトウェアとして，Just Right!6 Pro（ジャストシステム），一太郎 Pro4（ジャストシステム），PressTerm（NTT データ東北），Tomarigi（青山学院），Word（マイクロソフト）を比較し，Just Right!6 Pro を使用した。この方法を用いて評価ができるのは，文法や語句の使用法の適切性（誤字脱字，同音語誤り，同一助詞の連続など），文章全体や一文あたりの長さ（総文字数，文数，平均文長）である。これらの解析結果を前述の 3 項目に当てはめて評価を行った。

文章の自動評価手法の評価は，人間の評価と相関するかによって確かめられることが多いが，一方で本研究のように主観を排除した文章の評価手法の評価とは必ずしもよい相関を示すとは限らない。そこで，3 項目について人の評価との比較をすることに加えて，よい評価手法ではスコアに適度なばらつきが生じなければならないと仮定してスコアのばらつきによって本研究による方法の妥当性を考察した。人による評価は，3 項目のそれぞれについて 5 段階で評価した（0～4）。ソフトウェアによる「簡潔である」の評価においては，生スコアを平均値で除算することによって正規化した。

C．結果

AoMRC の分類では 82 の退院サマリに記載する項目が 23 に分類されている。記載すべき項目として，AoMRC 以外の文献から分類および項目を追加して 91 項目とし，そのうち 5 報以上の文献に登場した 38 項目を得た（この詳細は第 38 回医療情報学

連合大会の抄録参照)。定性的な要求事項として抽出したものを類似の要求事項をカテゴリにまとめて分類し，下記の 4 分類 9 項目となった。

【分類 1：完全性・正確性】内容に不足がない，診断などに関連のないことを記載していない，不適切なコピー＆ペーストをしていない，書かれている情報は正確である
【分類２：見読性・理解容易性】構造化されていて必要な情報にすぐにたどり着ける，文法が正しい
【分類３：言葉の使用】言葉遣いが適切である，診断名などは正確に書き，略語を使っていない
【分類４：分量・文字数】簡潔である

人による定性的な要求事項の評価は次の通りとなった。50 報の退院サマリのいずれもよく書けており，高いスコアになった（0 がよい評価）。

【文法が正しい】平均：0.14，SD：0.40，最低：0，最高：2
【言葉遣いが適切である】平均：0.24，SD：0.48，最低：0，最高：2
【簡潔である】平均：1.02，SD：0.89，最低：0，最高：3

次に，ソフトウェアを用いた評価は次の通りとなった。

【文法が正しい】平均：0.06，SD：0.24，最低：0，最高：1
【言葉遣いが適切である】平均：92.74，SD：19.92，最低：45，最高：135
【簡潔である】平均：5.00，SD：0.58，最低：3.9，最高：6.7

相関係数は，文法が正しい：0.12，言葉遣いが適切である：-0.08，簡潔である：0.14 となり，ほぼ相関がなかった。

なお，ソフトウェアによる評価の詳細は下記の通りであった。

【文法が正しい】修飾関係：0.06（SD: 0.24，Min: 0，Max: 1），並列関係：0.00（SD: 0.00，Min: 0，Max: 0），ら抜き表現：0（SD: 0，Min: 0，Max: 0），さ入れ表現：0（SD: 0，Min: 0，Max: 0），二重敬語：0（SD: 0，Min: 0，Max: 0），たりの脱落：0（SD: 0，Min: 0，Max: 0），べき止め：0（SD: 0，Min: 0，Max: 0）
【言葉遣いが適切である】誤字脱字：5.18（SD: 2.95，Min: 1，Max: 12），同音語誤り：0.08（SD: 0.27，Min: 0，Max: 1），送り仮名：0.2（SD: 0.49，Min: 0，Max: 2），漢字基準（常用漢字）：18.16（SD: 6.61，Min: 4，Max: 37），公用文：2.46（SD: 1.54，Min: 1，Max: 6），スペルチェック：63.56（SD: 20.80，Min: 15，Max: 105），表記ゆれ：2.68（SD: 1.65，Min: 0，Max: 7），同一助詞の連続：0.20（SD: 0.45，Min: 0，Max: 2）
【簡潔である】総文字数（空白除く）：2246.96（SD: 388.76，Min: 1,420，Max: 3,125），文数：85.44（SD: 33.39，Min: 21，Max: 169），平均文長：30.2（SD: 14.04，Min: 10，Max: 95）

D．考察

退院サマリに記載すべき項目として，投薬された薬剤名，退院後の診療行為，治療中の疾患，検査結果，入院中の治療，入院中の経過などが多くの文献で記載すべき項目として挙げられていた。退院時の担当医のように一部の項目は病院情報システムに登録された情報をそのまま引用できるが，38 項目のうち 24 項目（63%）はカルテの構造化データから自動で抽出することが難

しいと思われる項目であった。つまり，文章として記載をすることが求められる項目である。退院サマリの自動生成をする際には，これらの項目は自然言語処理を用いてカルテの自由記述から抽出し，要約をすることが求められる項目である。

記載すべき項目を文献から抽出する方法として，より多くの文献で触れられている項目を採用するという方法をとったが，この方法の課題に，特定の診療科や分野において重要と思われる項目が拾われないことがある。たとえば，褥瘡ケアのために入院前の ADL の把握が必要，終末期の慢性胃炎や長期間の介護を受けていた患者では栄養状態や緩和ケアの状況が必要，といった項目が例として挙げられる。対象領域を限定した研究を実施する際にはこの点を考慮し，当該領域の文献のみで分析を実施することが望ましい。

定性的な要求事項の 9 項目の中には，医療の知識がないと判断が難しいと考えられる事項が一部にあった。内容に不足がない，診断などに関連のないことを記載していない，不適切なコピー＆ペーストをしていないなど，9 項目のうち 5 項目が該当した。残りは知識ではなく形式から判断ができる可能性があるものであり，これらが自動採点の対象となり得ると考えられた。

定性的な要求事項の一部をスコア化することを試みた結果として，スコアを出すことはできるものの，人による評価との相関はなかった。相関なかった要因としてはいくつか考えられる。まず，本研究で用いた退院サマリはダミーとして医師に書き起こしてもらったものであるため，いずれもよく書けており，スコアの分散が小さくなるという点である。人による主観的な評価において退院サマリ間の差がほとんどなく，高いスコアに集中した。この傾向はソフトウェアによる評価において文法のエラーがほとんど指摘されていないことにも表れている。それ以外の指標に関するソフトウェアによる評価は機械的になされるため細かい差がついた。それなりの長さがある文章において文法のエラーを人が 5 段階で点数付けするというタスクは実際に行ってみると安定して実施をすることに困難が感じられ，この点はソフトウェアで評価を実施することの意義を示していると考えられる。

9 項目のうち 1 項目「診断名などは正確に書き，略語を使っていない」は，今回は評価対象に含めなかったが，医療用語辞書（病名，検査名，薬剤名など）を用いて解析をすることで対応が可能と思われる。なお，今回の解析において漢字の使用が不適切である（常用漢字ではない）との判定が頻発しており，その多くは医療用語によるものであった。また，校正ソフトウェアではスペルチェックを行うようになっており，非常に多くのスペルミスが指摘されているが（退院サマリあたり平均 63.6 箇所），検査名称などがスペルミスとして検出されており，これらも医療用語辞書を用いて対象から外す必要がある。

本研究の方法を実際の評価に使用する際には，9 項目のうち 1 項目「簡潔である」の判断のためにどの程度の数値であれば「簡潔である」であるかを事前に決める必要がある。このために，既存の「わかりやすい日本語」に関する研究の成果を援用することが可能と思われる。たとえば，一文あたりの文字数として 50〜80 文字と言われていたり，また外国人にわかりやすいのは平均 24 文字などと言われたりしている。なお，退院サマリは通常いくつかのブロックに分かれており，文が極端に短い箇所（単語の羅列）と長い箇所（経過に関する記述）とで一文あたりの文字数は極端に異なるので，今回の解析のように全体に 1 つの点数を付けることは不適切と考えられた。今後はこれらを分けて解析をする必要がある。

文章全体の文字数について参考になる数

値として，日本語の論文誌の抄録は 300～500 文字が上限として設定されていることが多い。この基準を踏まえると，本研究で用いた退院サマリは各文は短くまとまっているが，文の数はいずれも非常に多いと判定された。ただし，こうした適切さの基準はそれぞれ場面を限定して検討されていることもあり，つまり退院サマリとしてどれくらいの値が適切かはこれらとは一致しない可能性もある。よって理想的には，一般的な指標を参考にしつつ退院サマリとして適切な閾値を設定することが望ましい。

本研究の方法では 9 項目のうち 3 項目のみを評価可能と判断した。残りの 6 項目のうち 5 項目「内容に不足がない」,「診断などに関連のないことを記載していない」,「不適切なコピー＆ペーストをしていない」,「書かれている情報は正確である」,「構造化されていて必要な情報にすぐにたどり着ける」を判断するには医療の知識・経験およびそれに基づく内容の理解が必要であり，自動的に判定をするためには本研究とは大きく異なるアプローチが必要である。また，1 項目「構造化されていて必要な情報にすぐにたどり着ける」は医学的な知識は必要ないものの，一文を越えた範囲を解析する必要があるため本研究で使用したソフトウェアでは対応ができなかった。調査をした多くの文献で退院サマリの雛形を使用することの重要性が説かれていたが，雛形において記載箇所が複数に分けられている場合には，それぞれに記載されている内容を判定することで定量化が可能になると考えられる。先の 5 項目とは異なり，医学的な知識や内容の理解がなくとも，使用されている用語の頻度などで一定の判定は可能と想定される。

E．さいごに

本研究では，文献から抽出して整理した要件に基づいて退院サマリを評価するための方法を構築することを試みた。この方法で十分な水準で退院サマリを評価し得るものかは本研究によっては結論が出せなかったものの，退院サマリの自動評価方法の確立のために解決すべき様々な課題を明らかにすることができた。

F．研究発表

1．論文発表

なし

2．学会発表

森田瑞樹，奥村貴史，狩野芳伸，堀口裕正．退院サマリの自由記載は何を書くことが望ましいのか：文献レビュー，第38回医療情報学連合大会，2018年11月22～25日，福岡.