一般口演

一般口演21
医療データ分析９（機械学習・テキストマイニング２）

2018年11月25日(日) 10:40 ～ 12:10 D会場 (4F 413+414)

[4-D-2-5] 退院サマリー監査を支援する DPC判定アプリケーション

○鈴木 隆弘[1], 土井 俊祐[2], 木村 隆[1], 島井 健一郎[1], 嶋田 元[3], 畠山 豊[4], 本多 正幸[5], 松村 泰志[6], 横井 英人[7] （1.千葉大学医学部附属病院企画情報部, 2.東京大学医学部附属病院 企画情報運営部, 3.聖路加国際病院医療情報センター, 4.高知大学医学部附属医学情報センター, 5.長崎大学病院医療情報部, 6.大阪大学大学院医学系研究科医療情報学, 7.香川大学医学部附属病院医療情報部）

【目的】ＤＰＣコードを正しく付与することは病院にとって必須の課題であり、診療録監査の重要なポイントである。しかしながら限られたマンパワーの中で充分なチェックを行うことは難しい。本研究では退院サマリー監査のサポートとして、指定した期間の退院サマリーに対して一括処理を行い DPCコードを付与するアプリケーションを開発して、千葉大学において評価を行ったので報告する。【対象と方法】本アプリケーションは電子カルテからテキスト形式で出力された退院サマリーを読み込んでバッチ処理を行い、ベクトル空間モデルを応用してテキストデータベースとの照合を行い適切な DPCコードを付与して出力する。モデルデータとしては以前に報告した Web版類似症例検索と共通の、当院をはじめとした７病院の退院サマリー約３２万件を用いた。コーディング精度の評価として、診療情報管理士が自動判定された DPCコードと元のカルテを比較し、５（完全に一致）～１（見当はずれ)の５段階評価での判定を行った。【結果】投稿時点では予備段階として100件の退院サマリーを解析し、評価は平均で4.1点であった。今後は監査業務に応用しての実用性などを検討していく。本システムが順調に稼働すれば、医師はチェック結果を自己点検に用いることが出来るし、診療情報管理士は対象症例の絞り込みや監査の際の参考材料に利用することが出来る。

©一般社団法人 日本医療情報学会

退院サマリー監査を支援するDPC判定アプリケーション

鈴木隆弘[*1]、土井俊祐[*2]、木村隆[*1]、嶋田元[*3]、畠山豊[*4]、本多正幸[*5]、松村泰志[*6]、横井英人[*7]、島井健一郎[*1]
*1 千葉大学医学部附属病院企画情報部、*2 東京大学医学部附属病院企画情報運営部、
*3 聖路加国際病院医療情報センター、*4 高知大学医学部附属医学情報センター、*5 長崎大学病院医療情報部、
*6 大阪大学大学院医学系研究科医療情報学、*7 香川大学医学部附属病院医療情報部

DPC judgment application to support audit for discharge summary

Suzuki Takahiro[*1], Doi Shunsuke[*2], Kimura Takashi[*1], Shimada Gen[*3], Hatakeyama Yutaka[*4], Honda Masayuki[*5], Matsumura Yasushi[*6], Yokoi Hideto[*7], Shimai Ken-ichiro[*1]
*1 Department of Medical Informatics and Management, Chiba University Hospital, *2 Department of Healthcare Information Management, The University of Tokyo Hospital, *3 St. Luke's International Hospital Medical Information System Center, *4 Center of Medical Information Science, Kochi University, *5 Department of medical informatics, Nagasaki University Hospital, *6 Medical Informatics, Osaka University Graduate School of Medicine, *7 Department of Medical Informatics, Kagawa University Hospital

Abstract.
We developed an application that processes discharged summaries for a specified period and gives DPC code for the purpose of supporting summary audit, and made an evaluation at Chiba University hospital. This application reads the summary outputted in text from the electronic medical record and carries out batch processing, and verifies with the text database by applying the vector space model, gives an appropriate DPC code and outputs it as a list. We used approximately 320000 summaries as model data from 6 hospitals, Chiba University Hospital, Nagasaki University Hospital, Kagawa University Hospital, Osaka University Hospital, Kochi University Hospital and St. Luke's International Hospital. As a result of automatic DPC cross match judgment, the integrated model and its own model data showed high coincidence rate, and data of other hospitals was a slightly lower rate than that. The health information manager evaluated the coding accuracy of the application using 100 cases. Although it is difficult to use for audit directly, it was evaluated that it can be used as a pickup of the cautionary case and reference material. We will consider the practical utility when applied to audit work in the future.

Keywords: Textmining, Discharge Summary, DPC, Multi Center Study.

1. 背景

DPCコードを正しく付与することは病院にとって必須の課題であり、診療録監査の重要なポイントである。しかしながら限られたマンパワーの中で充分なチェックを行うことは難しい。我々は以前の研究[1)]で退院サマリーの文章からDPCコードを推定出来ることを示した。本研究では退院サマリー監査のサポートを目的に、指定した期間の退院サマリーに対して一括でDPCコードを付与する(処理を行う)アプリケーションを開発して、千葉大学において評価を行ったので報告する。

2. 対象

千葉大学病院、長崎大学病院、香川大学病院、大阪大学病院、高知大学病院、佐賀大学病院、聖路加国際病院の電子カルテに対応した退院サマリー収集と解析のためのプログラムを作成し、2012-2015年の4年間に入院した症例を対象として、退院サマリーと14桁のDPC情報の収集を行い、解析のためのモデルデータを作製した。対象となったサマリー数は千葉大学67266例、大阪大学58726例、香川大学20836例、高知大学 34994 例、長崎大学 61582 例、聖路加国際病院73065例で、総計は316459例であった。データの収集と基本的な処理は各施設内で行い、形態素分解後の連結不可能匿名化した情報のみを集約した。

Webアプリケーション検証用のデータとしては新たに、2018年8月1日から3日の間に千葉大学医学部附属病院を退院した患者から 100 例を無作為に選択し、退院サマリーを抽出して使用した。

3. 方法

3. 1. 形態素解析とベクトル化

抽出したサマリーから重要語を抽出してテキストデータベースを作成した。形態素解析には MeCab を使用し、サマリーの長さ、使われている単語、DPC との関連等について施設間の比較を行った。解析用辞書としては、我々がこれまでに作成した23万語の医療用語辞書とパラメディカル用医療辞書の ComeJisyo を併用した。得られた索引語の重み付けには tf×idf 法を用いてベクトル空間を構築した。

3. 2. DPC 自動判定と病院間クロスマッチ

各施設から症例数が 20 以上の DPC コードを抽出し、7 対 3 の比率で無作為にモデル作成用と検証用のサマリーに分け、検索対象のサマリーベクトル空間と入力されたサマリーを比較して最も類似したベクトルを検索し、DPC を判定した。類似度の算出はベクトルの内積によって行い、参加 6 病院各々の検証用データとモデルデータの組み合わせによる 6 対 6 のクロスマッチ判定および、全病院のデータを統合したモデルデータとの間での自動判定を、DPCコードの 14 桁全てと上位 6 桁のそれぞれで行った。

3. 3. Web アプリケーション

退院日を期間指定し、類似度の高い DPC コードを一覧表示することが出来る Web アプリケーションを、M言語データベースの Cache に装備されている CSP（Cache Server Page）を

用いて開発した。各施設別と全国共通のベクトルを切り替えて DPC を付与することが出来る機能を備えている(図1)。

図 1.　DPC 自動判定システム結果表示画面

アプリケーションの精度と使い勝手の検証のために、診療情報管理士と医事課入院請求担当職員による評価を 100 例に対して行った。疾患領域としては消化器疾患が多いものの、全ての MDC が含まれていた

4. 結果

4. 1. 退院サマリーの施設間比較

索引語の総出現回数は 1891525053 回で、単語数はそれぞれ、千葉大学 63441 語、大阪大学 88750 語、香川大学 75757 語、高知大学 50145 語、長崎大学は 78490 語で全体では 348912 語であった。1 サマリー内の平均単語数は千葉大 150 語、高知大 241 語、長崎大 68 語、大阪大 295 語、聖路加国際病院 210 語、香川大 242 語で、長崎大学が大きく短くなった他は前回の調査と同様の傾向を示した。

疾患とサマリーの長さとの関係は以前の報告と同様に、全ての病院で眼科疾患、耳鼻科疾患は短く、精神科疾患では長い傾向が認められた。

4. 2. DPC クロスマッチ判定

表 1 に疾患で決定される DPC6 桁までを用いたクロスマッチ照合の結果を示す。いずれの病院でも自施設のモデルデータと検証データの組み合わせでは、67～82%が一致と高い判定率を示した。異施設との間では 43～75%の判定率に低下するものの、全体としては 66%と高い判定率を示した。全施設を統合したモデルデータとの検証では 65～81%と自施設と同等の判定率を示した

表 1.　DPC6 桁によるクロスマッチ結果

モデル／検証	千葉大学	香川大学	高知大学	長崎大学	大阪大学	聖路加	統合データ
千葉大学	79%	58%	68%	60%	68%	60%	79%
香川大学	43%	67%	59%	40%	59%	43%	43%
高知大学	54%	67%	80%	71%	75%	69%	54%
長崎大学	50%	54%	65%	75%	66%	57%	50%
大阪大学	64%	65%	74%	63%	82%	67%	64%
聖路加	64%	57%	64%	60%	69%	76%	64%

表2に DPC14 桁全て一致での判定結果を示す。判定率は自施設同士で 55～72%、異施設との間でも 33～60%と 6 桁よりは下がるものの全体としての判定率は 54%で、統合モデルデータとの照合では、6 桁の場合と同様に自施設と同等の判定率を示した。

表 2.　DPC14 桁によるクロスマッチ結果《》800

モデル／検証	千葉大学	香川大学	高知大学	長崎大学	大阪大学	聖路加	統合データ
千葉大学	68%	39%	53%	47%	54%	46%	67%
香川大学	33%	55%	41%	29%	39%	31%	52%
高知大学	43%	47%	69%	57%	60%	53%	70%
長崎大学	43%	41%	55%	67%	55%	47%	68%
大阪大学	51%	44%	59%	51%	72%	52%	70%
聖路加	52%	41%	52%	50%	55%	67%	65%

4. 3. Web アプリケーションの評価

医事課職員による評価では、最終的に医事請求を行った DPC との比較を行ったが、14 桁では 27%、6 桁では 56%が一致した。クロスマッチの結果よりも一致率が低いが、別の日のサマリーでは同等の一致率を示しており、理由を調査中である。一致しない例の傾向としては、手術の有無や重症度の判定を誤っている例が多かった。診療情報管理士がサマリーから推測した 6 桁の DPC と比較した結果では、59%が一致した。一致しなかった例としては入院の契機となった病名や、既往歴に含まれている病名から DPC コードを判定している例が多く、最終的に否定された病名に影響された例もあった。また、サマリーのみでは最も医療資源を投入した病名を決めがたい症例も 3 例含まれていた。

5. 考察

新たな症例で行った DPC クロスマッチ自動判定の結果では、以前と同様に自施設のモデルデータでは高く、他施設のデータでは 10～20%程度低下するものの、全施設統合データでは自施設と同程度の判定率を示した。

退院期間を指定して DPC の候補を一覧表示するアプリケーションでは、結果をそのまま監査に用いるのは厳しいものの、要注意症例のピックアップや参考材料に利用することが出来るとの評価であった。類似度だけでなく、診断群分類決定にあたって考慮したデータ(キーワードやカルテの記載等)が参照出来れば、医師に DPC コーディングの提案が出来るのではないかとの意見もあり、本システムが順調に稼働すれば、医師はチェック結果を自己点検に用いることが出来るし、診療情報管理士は対象症例の絞り込みや監査の際の参考材料に利用することが期待出来る。

今後は精度を高めるとともに、監査業務に応用しての実用性などを検討していく予定である。

参考文献

1)　鈴木隆弘，土井俊祐，藤田伸輔ら．多施設間の統合退院サマリデータベースの構築．第 32 回医療情報学連合大会論文集 2012; 280-281)