

疫学研究に関する情報公開

福島県立医科大学医学部疫学講座では、本学倫理委員会の承認を得て、下記の疫学研究を実施します。関係各位のご理解とご協力をお願い申し上げます。

平成28年6月 福島県立医科大学医学部疫学講座 主任教授 大平 哲也

【研究課題名】福島県脳卒中発症登録調査における電子カルテデータの自然言語処理による利用に関する研究

【研究対象及び選定期間】

対象者数は悉皆登録のため2013年1月1日以降に福島県立医科大学付属病院を受診し入院された患者様のうち、脳卒中(160 くも膜下出血、161 脳内出血、162 その他の非外傷性頭蓋内出血、163 脳梗塞、164 脳卒中、脳出血又は脳梗塞と明示されないもの、165 脳実質外動脈の閉塞及び狭窄、脳梗塞に至らなかったもの、166 脳動脈の閉塞及び狭窄、脳梗塞に至らなかったもの、167 その他の脳血管疾患、168 他に分類される疾患における脳血管障害、169 脳血管疾患の続発・後遺症、G46 脳血管疾患における脳の血管(性)症候群)としてDPCコードがコーディングされている患者様を対象としています。2013年1月から12月までの間では約700名を対象としています。また、人工知能技術による自然言語処理の汎用性を確認するために、他の協力医療機関においても同様の調査を実施します。

【研究期間】 平成28年6月～平成33年5月

【研究の意義・目的】

福島県においては従来より脳卒中(平成22年度死因別年齢調整別男女別死亡率・脳血管疾患で、男性8位、女性6位)、心筋梗塞(平成22年度死因別年齢調整別男女別死亡率・急性心筋梗塞で男女ともに1位)の死亡率が高いことが報告されています。さらに、震災及びその後の放射線事故の影響により避難を余儀なくされた方々も多く居られ、生活習慣の変化や心理社会的ストレスの増大から循環器疾患発症の増加が懸念されています。地域住民における生活習慣病の発症動向を明らかにする必要があり、この唯一の方法が「県民健康調査」と突合可能な地域住民全員を対象とした脳卒中発症の悉皆登録です。しかしながら、現在「県民健康調査」のデータと突合可能な発症登録は、がんにおける地域がん登録のみです。循環器疾患については、すでに福島県では2009年1月より、福島県立医科大学医学部循環器・血液内科学講座を事務局とした急性心筋梗塞発症登録調査が行われているが、脳卒中に関する発症登録調査は行われておらず、県民健康調査とのデータの突合もできません。福島県民は以前より脳卒中死亡率が他の都道府県に比べて高く、今後さらに発症率・死亡率が上昇することが懸念されており、その対策は急務であり、福島県脳卒中発症登録ネットワークを立ちあげ、協力医療機関における採録を進めています。福島県における脳卒中発症登録数は、脳卒中発症登録を悉皆的に行っている滋賀県の発症数から、人口や脳卒中での死亡率から試算した場合、1年間で約5500件の脳卒中発症が予想される。採録の作業は非常に煩雑で、人力に頼るところが大きく、採録担当者の技量などによりデータの精度が安定しないことや、一件当たりの採録時間も症例によって

は 30 分～90 分を要する。このことは、脳卒中発症登録を悉皆的に行うにあたっての非常に高いハードルとなっていることは否定できません。

本研究の目的は、協力医療機関での脳卒中発症情報採録にあたり、その作業の煩雑さや効率的で精度の高い採録を可能とする手段として、医療機関において電子カルテ上に存在する診療情報を匿名化した上でテキスト情報として提供を受け、人工知能技術による自然言語処理を行うことで、1) 登録対象か登録除外かを一次的に振り分けること。2) 脳卒中発症登録データセットへの採録データ取込を自動化すること。3) 採録作業の軽減と、登録データの精度向上をはかることを目的としました。

【研究の方法】

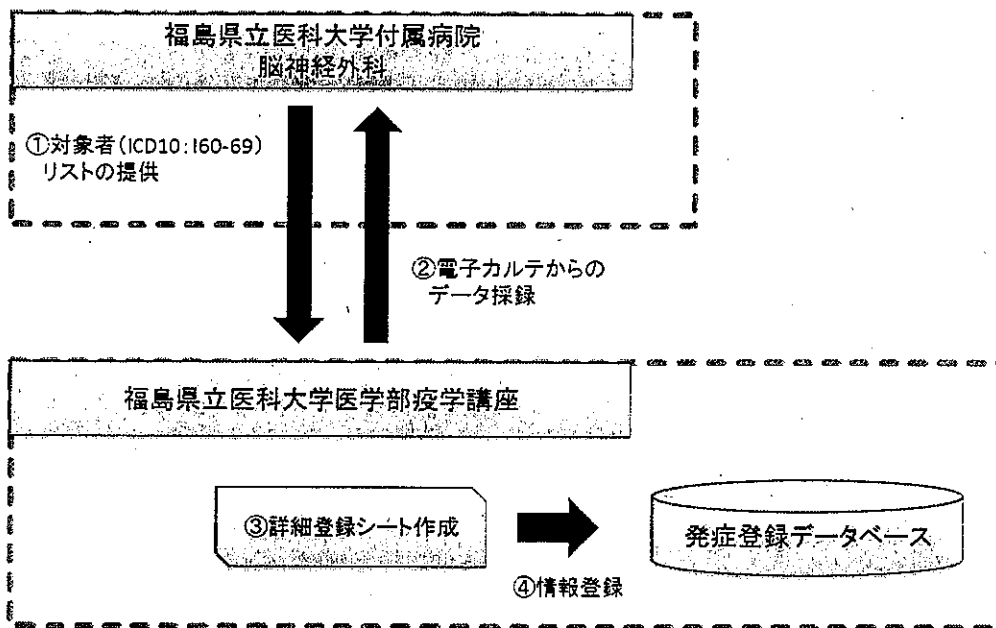


図 1. 現行の病院調査（附属病院）による電子カルテ採録の流れ

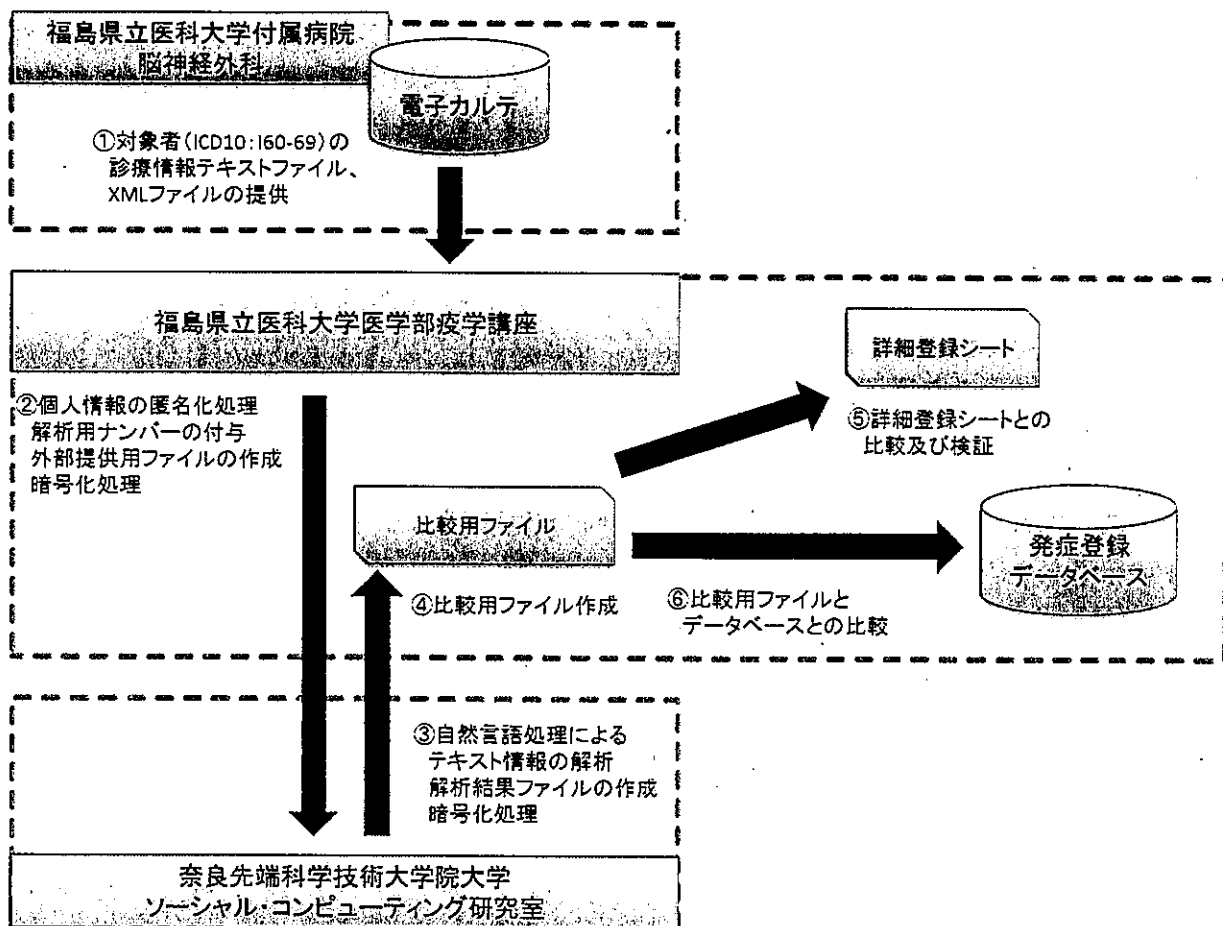


図 2. 本研究による自然言語処理により電子カルテデータ利用の流れ

図 1 に現在行われている病院調査（附属病院）と採録の手順について示します。脳神経外科から ICD-10 コードで脳卒中に関連のあるコードが付与された患者リストを元に、発症登録対象と考えられる患者の診療記録を電子カルテから閲覧し、詳細登録シートに転記した後に、登録対象となったものについて発症登録データベースに登録を行う手順です。この方法によるデメリットは、電子カルテの情報はデータの二次利用を目的とせず、脳卒中発症登録データセットに登録すべき項目について、電子カルテ上で記録のある場所を探し出し、目視により確認した後に紙の調査票への転記をするという煩雑な作業手順となっていることです。

そこで本研究では、図 2 に示すように、①脳神経外科から ICD-10 コードで脳卒中に関連のあるコードが付与された患者の電子カルテ上の診療記録をテキストファイルまたは XML ファイル形式で疫学講座が提供を受け、②発症登録にしか用いない個人同定用のコードを付与した上で、③個人属性情報を削除した上で、外部提供用のファイルを作成し、暗号化した後に共同研究機関である奈良先端科学技術大学院大学ソーシャルコンピューティング研究室へ電子媒体により送付します。奈良先端科学技術大学院大学ソーシャルコンピューティング研究室においては、人工知能を用いた自然言語処理により、発症登録に関する項目や治療方法、使用薬剤、病名及び検査記録の組み合わせから、当該患者が発症登録の対象か除外を確定するために必要となる解析結果用ファイルを作成し、暗号化した上で、疫学講座へ送付します。このとりやりとりされるファイルは、個人属性情報を含まず、元の電子カルテの内容も含まない解析結果のみのファイルであり、個人同定用のコードは、

外部へは提供しないことから、共同研究機関において個人が同定されたり、個人情報と紐づく形での診療情報を得ることはできない仕組みとなっています。疫学講座は共同研究機関から送付された解析結果ファイルを元に、比較用ファイルを作成し、詳細登録シートの記載内容及び、登録対象か除外かの結果を自然言語処理による結果と照合し、比較を行います。これにより、自然言語処理による判定基準の精度について検証を行う。既に登録されている発症登録データベースの内容についても、比較用ファイルとの検証を行い、判定基準の精度及び、自然言語処理による結果が、脳卒中発症登録データセットの必要な項目を満たしているかについて検証を行います。検証結果について疫学講座から共同研究機関へフィードバックを行うことで、自然言語処理用データベースへの知識情報の追加と処理プログラムの改良を加えることで、より精度の高い自然言語処理を行うことが可能となります。共同研究機関の責任として、匿名化され個人情報と紐づけられないことがない診療情報であっても、流出などがおきないように細心の注意を払い、情報保護と情報セキュリティの観点から、必要十分な対策をとることとします。

【研究組織、研究機関名】

研究責任者

医学部疫学講座 主任教授 大平哲也

主任研究者

放射線医学県民健康管理センター疫学健康調査支援部門 疫学室 中野裕紀

研究分担者

脳神経外科講座 教授 齋藤 清

脳神経外科講座 准教授 佐久間 潤

脳神経外科講座 講師 佐藤 拓

脳神経外科講座 助手 織田恵子

放射線医学県民健康管理センター健康調査支援部門疫学室 助手 弓屋 結

学外研究者

奈良先端科学技術大学院大学ソーシャル・コンピューティング研究室

特任准教授 荒牧英治

奈良先端科学技術大学院大学ソーシャル・コンピューティング研究室

博士研究員 若宮翔子

【研究者が保有する個人情報について】

研究者が保有する個人情報に関し、被験者ご本人又は代理人の方が開示、訂正、利用停止及び第三者への提供の停止等の請求を行う場合、「福島県個人情報保護条例」に基づく手続きが必要となります。なお、開示等を行う場合、請求者には文書等の交付に係る費用（コピー代等）をご負担いただきます。

【本研究に関する問合せ先】

○研究内容に関する問合せの窓口

〒960-1295 福島県福島市光が丘 1

公立大学法人福島県立医科大学医学部疫学講座 担当 大平哲也

電話:024-547-1738 FAX: 024-547-1244

E-mail:teohira@fmu.ac.jp

○個人情報に関する窓口

〒960-1295 福島県福島市光が丘1

公立大学法人福島県立医科大学 総務課 大学管理係

電話:024-547-1007 FAX :024-547-1995

○その他ご意見の窓口

〒960-1295 福島県福島市光が丘1

公立大学法人福島県立医科大学 医療研究推進課 研究支援担当

電話:024-547-1825 FAX: 024-547-1991

E-mail:rs@fmu.ac.jp