

一般口演 | 医療支援

一般口演2 医療支援

2021年11月19日(金) 09:10 ~ 11:10 F会場 (2号館2階224)

[2-F-1-04] 米国診療データを活用した2型糖尿病向け処方選択支援技術の開発と評価

*垂水 信二¹、竹内 渉¹、Chalkidis George¹、Loya Salvador²、桑田 純一¹、Flynn Michael³、Turner Kyle³、Sakaguchi Farrant³、Weir Charlene²、Taft Teresa²、Shields David²、Warner Phillip²、Kukhareva Polina²、伴 秀行¹、Kawamoto Kensaku²（1. 株式会社日立製作所 研究開発グループ, 2. Department of Biomedical Informatics, University of Utah, 3. University of Utah Health）

*Shinji Tarumi¹、Wataru Takeuchi¹、George Chalkidis¹、Salvador R. Loya²、Junichi Kuwata¹、Michael C. Flynn³、Kyle M. Turner³、Farrant H. Sakaguchi³、Charlene R. Weir²、Teresa Y. Taft²、David E. Shields²、Phillip B. Warner²、Polina V. Kukhareva²、Hideyuki Ban¹、Kensaku Kawamoto²（1. 株式会社日立製作所 研究開発グループ, 2. Department of Biomedical Informatics, University of Utah, 3. University of Utah Health）

キーワード：Clinical decision support system, Disease management, Artificial intelligence, Chronic disease

【背景と目的】2型糖尿病は代表的な生活習慣病のひとつであり、長期にわたる適切な治療を必要とする。患者最適な処方選択を支援するため、我々はこれまでに、治療種別のアウトカムを実臨床データから予測する処方選択支援システムの開発を進めてきた。本発表では、予測性能向上を目的に開発した、治療遷移構造に基づくアウトカム予測手法について報告する。【方法】米国ユタ大学の診療データ27904例を用いて、3か月後にHbA1c<7%などの治療目標を達成する確率を、治療種別に予測する技術を開発した。提案法は、データ内の処方パターンを遷移グラフでモデル化することで、治療状態に応じた予測と処方推薦を実現する。7割のデータで学習したモデルを3割のデータに適用し、以下を評価した。1) 従来法である勾配ブースティング決定木を比較対象に、治療達成確率の予測誤差を Brier Scoreで評価した。2) 米国糖尿病学会の診療指針を参考に、最大の確率で推薦された処方の特徴を調査した。また、予測の妥当性を臨床医3名に確認いただいた。本研究はユタ大学と(株)日立製作所研究開発グループで定める倫理審査基準に則り審査、承認され、実施された。【結果および考察】1) 従来法の予測誤差0.162を下回る0.158を得た。キャリブレーション誤差は、従来法0.013を大きく下回る0.002（84.6%減）を得た。2) 従来法の推薦処方では、指針で推奨される段階的に強度を増加する処方が57.8%含まれたが、提案法では98.6%含まれ、より指針に則していることが確認された。また、提案法の推薦処方が、臨床医の感覚と一致する旨を確認いただいた。以上より、従来技術に対する優位性と、臨床的有用可能性を確認した。【結論】患者最適な処方選択を支援する2型糖尿病向け処方選択支援技術を開発した。今後、本技術の臨床的有用性を前向き臨床研究にて評価する。

米国診療データを活用した2型糖尿病向け処方選択支援技術の開発と評価

垂水信二^{*1}、竹内渉^{*1}、Salvador R. Loya^{*2}、桑田純一^{*1}、George Chalkidis^{*1}、Michael C. Flynn^{*3}、Kyle M. Turner^{*3}、Farrant H. Sakaguchi^{*3}、Charlene R. Weir^{*2}、Teresa Y. Taft^{*2}、David E. Shields^{*2}、Phillip B. Warner^{*2}、Polina V. Kukhareva^{*2}、伴秀行^{*1}、Kensaku Kawamoto^{*2}

^{*1} 日立製作所研究開発グループ、^{*2} Dept. of Biomedical Informatics, Univ. of Utah、

^{*3} Univ. of Utah Health

Development and Evaluation of Pharmacotherapy Decision Support Technology for Type-2 Diabetes Mellitus using Clinical Data in the United States of America

Shinji Tarumi^{*1}、Wataru Takeuchi^{*1}、Salvador R. Loya^{*2}、Junichi Kuwata^{*1}、George Chalkidis^{*1}、Michael C. Flynn^{*3}、Kyle M. Turner^{*3}、Farrant H. Sakaguchi^{*3}、Charlene R. Weir^{*2}、Teresa Y. Taft^{*2}、David E. Shields^{*2}、Phillip B. Warner^{*2}、Polina V. Kukhareva^{*2}、Hideyuki Ban^{*1}、Kensaku Kawamoto^{*2}

^{*1} Research & development group, Hitachi Ltd., ^{*2} Dept. of Biomedical Informatics, Univ. of Utah、

^{*3} Univ. of Utah Health

Type-2 diabetes mellitus (T2DM) is a chronic disease requiring long-term treatment. The established treatment target for most patients is controlling hemoglobin A1c (HbA1c) levels to less than 7%; however, about half of patients fail to achieve this target. To help address this problem, the authors have developed a machine learning driven pharmacotherapy decision support system which can be integrated with the electronic health record (EHR) using standards-based clinical decision support (CDS) interoperability frameworks. A novel prediction algorithm, Treatment Pathway Based Estimation (TPGE), was proposed for the system to predict the likelihood of alternate treatment strategies achieving care goals (such as an HbA1c level below 7% within 3 months). The algorithm was evaluated using real world EHR datasets from University of Utah Health in i) treatment outcome prediction tasks and ii) treatment recommendation tasks. A conventional machine learning method, Gradient Boosting Tree (GBT), was used as a baseline method. The authors confirmed that i) TPGE outperformed GBT in reliability of prediction and ii) TPGE recommended treatment options that followed clinical guideline more accurately. The system was successfully integrated with the EHR at University of Utah Health.

Keywords: Clinical Decision Support, Machine Learning, Type-2 Diabetes

1. 緒論

2 型糖尿病は代表的な生活習慣病のひとつであり、長期にわたる治療の継続が必要とされる。米国の診療ガイドライン¹⁾では、ヘモグロビン A1c(HbA1c)を 7%未満に維持することが一般的な治療目標として定められているが、約半数の患者がこの目標を達成できていないとの報告もあり²⁾、患者状態に応じた最適な治療の実現は重要な課題である。

臨床現場において治療を効果的に支援する技術として、Clinical Decision Support (CDS) システムが注目されている。特に米国においては、国際 HL7 協会(Health Level Seven International)主導で開発が進められている FHIR(Fast Healthcare Interoperability Resources)³⁾や電子カルテとの認証やデータ取得方法を規定する SMART on FHIR⁴⁾フレームワークなど、医療情報システム間の相互運用性(Interoperability)を担保するための次世代標準規格の整備が進んでおり、電子カルテベンダーに依存しない CDS システムの柔軟な活用が加速している。

従来の CDS システムは、診療ガイドラインなど医学的コンセンサスに基づいたルールを用いて注意喚起や情報提示を行うものが主であったが、近年は、機械学習に基づく予測技術を活用した CDS システムの研究開発が進み、合併症の発症リスク予測などの活用事例が報告されている⁵⁾。一方、患者に最適な治療選択を支援可能な CDS システムは、未だ発展段階にある。特にデータ分析における重要な課題のひとつは、患者毎の治療結果を正確かつ頑健に予測することの困難さである。例えば決定木を基にした Random Forest (RF)⁶⁾や Gradient Boosting Tree (GBT)⁷⁾といった機械学習の手法

は、診断支援やリスク予測の事例で高い予測性能が確認されているものの、治療アウトカム予測においては、予測にバイアスが生じる懸念が報告されており⁸⁾、更なる技術開発が望まれている。加えて、一般的な機械学習の手法は、学習データに含まれるパターンを教師として学習するため、仮に学習データの特徴から乖離した患者を予測する場合、外挿による想定外の結果が生じる可能性があり、臨床で活用する障害が大きかった。

2. 目的

2 型糖尿病患者の最適な処方選択を支援するため、我々はこれまでに、治療種別のアウトカムを実臨床データから予測する処方選択支援システム⁹⁾の開発を米国にて進めてきた。開発に際しては、予測を電子カルテから活用可能にすること、開発成果の他院への展開性、を考慮し、米国 Utah 大学と共同で上述の FHIR などの次世代標準規格に準拠したアプリケーションとして実装を進めてきた。

本研究では、開発中の処方選択支援システムにおいて、より正確かつ頑健な治療アウトカム予測の実現を目的とし、新たな治療アウトカム予測技術を開発した。Utah 大学の関連クリニックの実診療データを用いて、その予測性能と推薦される処方の性質を評価し、手法の医学的有用性を検証した。

3. 方法

提案法は、治療手段の遷移構造を表現する治療遷移グラフを過去の診療データから学習しモデル化することで、患者毎の治療状態に応じた治療アウトカムの予測と処方推薦を実現する。以下、提案法の詳細と、性能評価の方法を説明す

る。

尚、本研究は Utah 大学および(株)日立製作所研究開発グループで定める倫理審査基準に則り審査、承認され、実施された。

3.1. 治療遷移グラフを用いたアウトカム予測法

一般に治療法を選択する場合、いくつかの候補から最適と期待されるものを選択する。この意思決定を効果的に支援するためには、候補となる治療手段のアウトカムを比較可能な形で予測できることが望ましい。具体的には、各予測が、背景が適切に調整された過去症例に基づき実施されることが望ましい。例えば、患者に対する第一選択薬の候補を複数検討している場合、それらの予測は、第一選択薬を処方した過去症例に基づくべきであり、既に投薬を実施中の患者に対する追加処方の症例は用いるべきではないと考えられる。

そこで提案法は、上述の治療遷移グラフを構築し、各治療状態において想定される治療遷移パターン毎に、実際にその治療遷移パターンを経験した症例に基づきアウトカムを予測することで、現状の治療状態に関する交絡の影響を低減する。以下、本手法のモデル構築プロセスを、1) データ整形処理、2) グラフ構築処理、3) モデル構築処理、の順に説明し、最後にモデルを用いた 4) アウトカム予測処理、について説明する。

3.1.1. データ整形処理

本処理では、蓄積された電子カルテ情報を、分析可能な表形式データに整形する。表の各行は、ある患者に対するひとつの投薬治療に対応し、現状の治療の情報、新たに実施した治療の情報、治療前後の HbA1c の値、その他関連する特徴量として、患者背景情報(性別、年齢)、治療前の検査値情報(中性脂肪、eGFR など8種類)、生体情報(体重、血圧など4種類)、合併症に関する診断情報(Clinical Classification Software に基づく135分類)の計149種の特徴量を含む。HbA1c の計測間隔は、一般の糖尿病患者のフォローアップが3か月単位で行われることに鑑みて90日±30日以内に計測された症例を抽出した。治療薬の分析単位は分析可能な症例数の想定から、薬剤分類(メトホルミンやDPP-4阻害剤などの分類7種類)でカテゴリ化したものを用いた。

3.1.2. グラフ構築処理

本処理では、整形した表形式データから、治療遷移グラフを構築する。治療遷移グラフは、複数のノードと有向エッジから構成される有向グラフである。ノードは、ひとつの治療オプション(例えば2剤併用処方の場合、2つの治療薬の組み合わせ)を表し、有向エッジは、治療オプション間の遷移構造を表す。図1は、治療遷移グラフの例を表す図である。例えば「投薬なし」ノードは、「投薬なし」ノード自身に遷移する有向エッジと、「ビグアナイド薬」に遷移する有向エッジの2つのエッジを有するが、これらはそれぞれ、「投薬なし状態の維持」、「ビグアナイド薬の新規処方の開始」、を意味する。治療遷移を表現する有向エッジは、整形した表形式データに含まれる症例の中で、当該の治療遷移を経験した症例を抽出し、その症例数が10以上の場合に生成した。各有向エッジは、それぞれ対応する治療遷移を経験した症例の治療前のHbA1cの分布から計算した、HbA1cに関する有効範囲を有する。これにより、現状の処方薬とHbA1cの情報に基づき、各治療遷移がその患者にとって適切か否かの判定が可能になる。

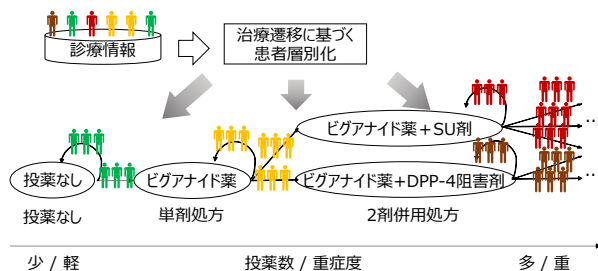


図1 治療遷移グラフ

3.1.3. モデル構築処理

本処理では、グラフの各有向エッジに対応するアウトカム予測モデルを構築する。モデルは各有向エッジに対応する症例情報から構築される。モデルの説明変数は整形した表形式データに含まれる、149種の特徴量を用いる。目的変数には、将来のHbA1cの値に関して設定した値を達成するかどうかの確率を用いる。例えば、3か月後にHbA1cが7%未満にコントロールできる確率、などを用いる。目標値は{7.0%, 7.5%, 8.0%, 8.5%}の4つから患者個別に選択することとする。

モデルには、症例数が少ない場合でも頑健な予測を実現するため、Support Vector Regression (SVR)¹⁰⁾とLogistic regression(LR)¹¹⁾を階層的に構成したものを用いた。モデル構造を図2に示す。SVRの層は、治療アウトカムに対して、最も支配的な予測因子である現在のHbA1cの値の影響を非線形にモデル化するために導入した。LRの層は、HbA1c以外の関連する予測因子の影響を考慮するため、および、出力値を確率値に変換するために導入した。

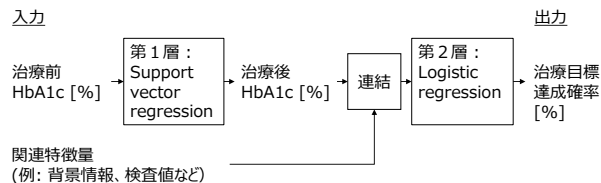


図2 モデル構造

3.1.4. アウトカム予測処理

本処理は、学習済みの治療遷移グラフおよび予測モデルを用いて、各治療候補のアウトカムを予測する処理である。まず患者の現在の治療状態に基づき、治療遷移グラフにおける患者のノード位置を同定する。次に、予測対象の治療オプションが、患者のノードと繋がっている有向エッジに含まれているかを判定する。含まれる場合、当該の有向エッジに対応する予測モデルを用いて治療アウトカムを予測する。含まれない場合、または、含まれていても、当該の有向エッジに対応するHbA1cの有効範囲から、予測対象の患者のHbA1cが乖離する場合は、予測を不適と判定し、予測を実施しない。この処理により、学習データから乖離する症例や予測パターンに対する外挿を抑制し、信頼性の低い情報に基づく支援を抑制する。

3.2. 評価

米国 Utah 大学の実診療データ 27904 例(6 年分 2012～2017 年)を用いて、以下の観点から提案法を評価した。

3.2.1. 精度評価

予測確率の精度を評価するため、確率に関する誤差指標である Brier Score¹²⁾を用いて検証した。患者が重複しないよう、データを7:3の割合で学習用セットと評価用セットに分割し、学習用セットで5分割交差検証を行った後、構築したモデルを評価用セットに適用し、治療アウトカムの予測精度を評価した。比較対象には、GBTを用いた。加えて、取得された予測精度の詳細な特徴を調査するため、Brier ScoreをReliability、Resolution、Uncertaintyの3つの要素¹³⁾に分解し、各要素の差を検証した。

3.2.2. 治療薬推薦シミュレーション評価

学習用セットで構築したモデルを用いて、最も治療アウトカムが高いと期待される治療手段を推薦するシミュレーションを行い、その結果を評価した。シミュレーション対象は評価用セットの中で、HbA1cが7%以上かつ、実際に1剤以上の治療薬を処方された3715例とし、処方数を維持もしくは増加させる治療を治療推薦の候補とした。GBTと提案法のそれぞれで推薦された治療手段に関して、診療ガイドライン⁹⁾との整合性を評価した。また、予測の妥当性を臨床医3名に確認いただいた。

4. 結果

提案法はGBTの予測誤差0.162を下回る0.158を得た。テストデータに対するBrier Scoreを分解した結果、Reliabilityは従来法0.013を大きく下回る0.002(84.6%減)であり、キャリブレーションの精度が大幅に向上したことが確認された。

従来法で推薦された処方では、診療ガイドラインで推奨される段階的に強度を増加する処方(1剤入替または1剤追加する処方)が57.8%含まれたが、提案法では98.6%含まれ、より指針に則する可能性が確認された。また、提案法の推薦処方が、臨床医の感覚と一致する旨を確認いただいた。以上より、従来技術に対する優位性と、臨床的有用可能性を確認した。

表1 予測性能評価結果

評価方法	Brier Score	
	従来法(GBT)	提案法
5分割交差検証	0.196±0.024	0.144±0.003
テストデータでの評価	0.162	0.158

表2 テストデータに対するBrier Scoreの分解結果

分解後の要素	従来法(GBT)	提案法
Reliability	0.013	0.002
Resolution	0.101	0.094
Uncertainty	0.250	0.250

表3 推薦された治療の追加薬剤数別の集計結果

追加薬剤数	推薦数 / 推薦割合			
	従来法(GBT)		提案法	
0	562	57.8%	1320	98.6%
1	1586		2344	
2	1377	42.2%	51	1.4%
3	190		0	
合計	3715	100.0%	3715	100.0%

5. 考察

本研究では、正確かつ頑健な治療アウトカム予測の実現を目的として、新たな治療アウトカム予測技術を提案し、米国の

実臨床データを用いて、その医学的有用性を評価した。

GBTを上回る予測精度を達成し、特に、キャリブレーション誤差が大きく改善した。この理由は、提案法は現状の治療状態に応じた層別化をすることで、予測対象の患者と類似する症例の情報だけを予測に活用できたためと考えられる。

推薦シミュレーションでは、追加処方数の少ない治療手段が多く推薦され、より効果的な治療に繋がる可能性が示唆された。この理由は、外挿を抑制できないGBTに対し、提案法は、過去の類似症例で実際に処方された治療薬に絞り込んだ推薦が可能であるためと考えられる。なお、本シミュレーションはHbA1cの目標達成のみに着目し、低血糖リスクや合併症発生病リスクなどの要素は考慮できていないため、より現実に応じた条件下での追加評価が将来の課題である。

また、本研究の評価結果はいずれも後ろ向き解析の結果であるため、実臨床上の有用性の評価には、前向きの臨床研究が必要である。さらに、本研究で用いたデータはUtah大学のデータに限定されているため、他の医療機関に対する一般化可能性の検証も併せて進める必要がある。

6. 結論

患者最適な処方選択を支援する2型糖尿病向け処方選択支援技術を開発した。今後、本技術の臨床的有用性を前向き臨床研究にて評価する。

参考文献

- 1) American Diabetes Association. Standards of medical care in diabetes—2018. Diabetes Care 2018;41(Suppl 1):S1–S2.
- 2) Stark Casagrande S, Fradkin JE, Saydah SH, Rust KF, Cowie CC. The prevalence of meeting A1C, blood pressure, and LDL goals among people with diabetes, 1988–2010. Diabetes Care. 2013;36(8):2271–9.
- 3) HL7.org. HL7 FHIR. [https://www.hl7.org/fhir/ (cited 2021-Aug-18)].
- 4) Mandl KD, Mandel JC, Murphy SN, et al. The SMART Platform: early experience enabling substitutable applications for electronic health records. J Am Med Inform Assoc. 2012;19(4):597–603.
- 5) Dagliati A, Sacchi L, Tibollo V, et al. A dashboard-based system for supporting diabetes care. J Am Med Inform Assoc 2018;25(05):538–547.
- 6) Breiman L. Random forests. Mach Learn 2001;45(01):5–32.
- 7) Friedman J. Greedy function approximation: a gradient boosting machine. Ann Stat 2000;29:1189–1232.
- 8) Fang G, Annis IE, Elston-Lafata J, Cykert S. Applying machine learning to predict real-world individual treatment effects: insights from a virtual patient cohort. J Am Med Inform Assoc 2019;26(10):977–988.
- 9) Tarumi S, Takeuchi W, Chalkidis G, et al. Leveraging Artificial Intelligence to Improve Chronic Disease Care: Methods and Application to Pharmacotherapy Decision Support for Type-2 Diabetes Mellitus. Methods Inf Med. 2021;60(S 01):e32–e43.
- 10) Drucker H, Burges CJC, Kaufman L, Smola A, Vapnik V. Support vector regression machines. Adv Neural Inf Process Syst 1997;9:155–161.
- 11) Cox DR. The regression analysis of binary sequences (with discussion). J R Stat Soc B 1958;20:215–24.
- 12) Brier GW. Verification of forecasts expressed in terms of probability. Mon Weather Rev 1950;78:1–3.
- 13) Murphy, AH. A new vector partition of the probability score. J. Appl. Meteorol. Climatol 1973;12(4):595–600.