

学位論文題名

Representing and Processing Medical Knowledge
for EBM Support in HIS

(医学知識の表現と処理に関する基礎研究)

—HIS における EBM 支援を目標として)

学位論文内容の要旨

近年、情報技術 (IT) の医療への応用は日常の臨床業務にまで拡大し、着実にその応用範囲は広がっている。日本では 1990 年代初期に導入が始まったオーダリング・システムを中心とする病院情報システム (HIS) の普及により、病院での情報の管理が開始されたと言っても過言ではない。現在では、電子カルテの急速な普及により HIS の利用範囲が拡大しており、根拠に基づく医療 (evidence-based medicine: EBM) を実践する上でも電子カルテの開発と利用が課題となっている。更に、これらの情報技術 (IT) を利用することにより、医療の費用と質の管理を行うことが可能になった。

臨床医の情報ニーズを理解することは、臨床情報システムの開発と利用に不可欠である。特に、近年の EBM の概念の普及に伴い、臨床医が必要とする最新の医学的な根拠を提供するための意思決定支援システムの開発が医療情報学分野で注目されている。知識工学の一つである知識ベースシステムを応用した医学概念のコンピュータ生成手法及び医学概念に関連づけられた分析手法を用いることにより、臨床医に不必要な情報を削減し、必要とする情報を的確に取得することを可能となる。この研究を進めるためには、標準化された医学用語集が医学知識表現を表現する上で不可欠である。標準化された用語集を作成するためには、含まれる概念、用語の分類と階層の枠組み、提供する用語などが要求されている範囲や質や満たしているかを検討しなければならない。知識ベースシステムにおける知識管理は、オントロジー (ontology) と言われる知識の獲得、分類、構造、配信の各過程が必要となる。オントロジーは、「ある概念を様々な分野を対象として横断的に再利用するために考慮された記述」や、「特定分野の知識のモデル化、すなわち、概念、属性、概念相互の関係」と定義され、用語集が単にある領域で使用する単語に限定されているのに対して、オントロジーはある領域における関連した概念とその関係を表している。近年、医学領域のオントロジーの構築が医療情報学においても注目されてきた。明確化された医学概念 (医学オントロジー) を計算可能なモデルに導入することにより、医学領域での知識表現システムの構築が可能となり、英語においてはすでに構築が試みられている。しかしながら、医学オントロジーの構築は、言語構造に依存するため複雑な解析と時間を必要とし、日本語での医学オントロジーの構築は全くなされていないのが現状である。

本論文では、HIS における EBM 支援を目的として、文脈に基づくオントロジーの構築と標準化された臨床用語の評価を通じて、臨床領域における医学知識の表現と処理の必須要件の役割について検討するとともに、臨床医が EBM を実践するために必要とする情報ニーズの調査を行った。

第 1 章の概説のあと、第 2 章においては、自然言語処理 (natural language processing: NLP) を導入した形式概念分析 (formal concept analysis: FCA) モデルによるオントロジー構築支援システムの構築について記述した。システムのユーザ・インタフェースは、Protege-2000 の JAVA タブ・プラグインを使用し、368 編の退院サマリーをもとに、オントロジー構築

を目的として、日本語標準病名集 (MEDIS ver2.0) を参照しての分析を行った。その結果、抽出された病名複合語の 73% が臨床の視点から医学概念として十分に意味のあるものであり、抽出された医学概念ペアの 57.7% は、臨床の視点から妥当であると判定され、医学の概念抽出に関して高い精度があることが示された。FCA を用いたこのオントロジー構築支援システムにより、専門医に対し多くの言語情報と文脈に基づく知識獲得の支援をすることが可能となった。

既存の標準病名集の評価・分析も、今後の標準病名集の開発方針を決定する上で不可欠である。米国国立医学図書館 (NLM) によって作成された統合医学用語システム (UMLS) では、意味毎における用語の再分類を行うことによって、医学用語の概念の類型を示すことが可能としている。第 3 章では、UMLS を用いた意味に基づく概念類型の分類、及び、既存の英語の臨床用語集と比較することにより、日本語標準病名集の評価を行った。その結果、日本語標準病名集は ICD10 と同じ概念類型の分類であることが明らかになり、今後、使用者側の概念類型を取り入れることが病名集の開発に重要であることが示唆された。また、UMLS による高度な意味に基づく概念類型が既存の日本語及び英語での臨床用語集の理解を促すと共に、その範囲を評価することに役立つモデルとなることが明らかになった。

第 4 章では、臨床研究におけるコンピュータと HIS の利用に関する日本と中国の比較、及び日本の臨床医を対象とした HIS 構築に不可欠な診療における詳細な情報ニーズの調査を目的として、行った調査結果をまとめた。質問の内容は、コンピュータの読み書きの能力、臨床診療のためコンピュータに基づく検索、インターネットの使用、HIS 上情報の使用、患者と電子メール連絡の状況とし、調査対象は、中国の 42 大病院の臨床医 172 名、日本の 72 大学病院の臨床医 568 名とした (回収率は平均 73.4%)。その結果、日本の臨床医は臨床研究のためにコンピュータ、e-mail、文献データベースの使用頻度が中国よりも高いが、HIS を利用した臨床データの収集では、両国の利用頻度に有意差はないことが明らかになった。

第 5 章では、日本における根拠に基づく医療 (EBM) に対する臨床医の意識と問題点、EBM を実践するために必要な外来診療時の臨床医の情報ニーズを明らかにするために、北海道大学病院に勤務する 100 名の臨床医を対象とした調査結果をまとめた。質問の内容は、臨床問題の種類と発生頻度、問題解決のための情報源の種類と使用頻度、EBM を実践するための HIS に期待する情報支援機能、とした。その結果、日本の多くの臨床医は根拠に基づく医療 (EBM) の実践を試みているが、EBM のための情報学的手法が十分に普及していないことが明らかになった。

これらの研究より、臨床医が EBM を実践するために必要とする情報ニーズを明らかにするとともに、日本語での医学オントロジーによる臨床知識の表現と処理を通じて、今後の EBM 支援を可能とする病院情報システム開発に関する研究の基礎を構築した。

学位論文審査の要旨

主 査 教 授 櫻 井 恒太郎

副 査 教 授 北 島 顕

副 査 教 授 寺 沢 浩 一

学 位 論 文 題 名

Representing and Processing Medical Knowledge for EBM Support in HIS

(医学知識の表現と処理に関する基礎研究

—HIS における EBM 支援を目標として)

本論文は、HIS における EBM 支援を目的として過去4年間に行った4つの研究(何れも英文論文および国際学会で公表)を集成したもので、医療文書に基づくオントロジーの構築、標準化された臨床用語集の評価による臨床領域における医学知識の表現と処理の必須要件の検討を行うとともに、臨床医が EBM を実践するために必要とする情報ニーズの調査を合わせて行った。

第1章は概説である。第2章では自然言語処理(Natural Language Processing: NLP)を導入した形式概念分析(Formal Concept Analysis: FCA)モデルによるオントロジー構築支援システムの構築について記述した。システムのユーザ・インタフェースは、Protege-2000 の JAVA タブ・プラグインを使用し、368 編の退院サマリーをもとに、オントロジー構築を目的として、日本語標準病名集 (MEDIS ver2.0) を参照しての分析を行った。その結果、抽出された病名複合語の 73%が臨床の視点から医学概念として十分に意味のあるものであり、抽出された医学概念ペアの 57.7%は、臨床の視点から妥当であると判定され、医学の概念抽出に関して実用的な精度を有することが示された。FCA を用いたこのオントロジー構築支援システムにより、専門医に対し多くの言語情報と文脈に基づく知識獲得の支援をすることが可能となった。

第3章では、統一医学用語システム(Unified Medical Language System: UMLS:)を用いた意味に基づく概念類型の分類、及び、既存の英語の臨床用語集と比較することにより、日本語標準病名集の評価を行った。その結果、日本語標準病名集は ICD10 と同じ概念類型の分類であることが明らかになり、今後、使用者側の概念類型を取り入れることが病名集の開発に重要であることが示唆された。また、UMLS による高度な意味に基づく概念類型が既存の日本語及び英語での臨床用語集の理解を促すと共に、その範囲を評価することに役立つモデルとなることが明らかになった。

第4章では、臨床研究におけるコンピュータと病院情報システム(Hospital Information System: HIS)の利用に関する日本と中国の比較、及び日本の臨床医を対象とした HIS 構築に不可欠な診療における詳細な情報ニーズの調査を目的として、調査結果をまとめた。その結果、日本の臨床医は臨床研究において e-メールおよび文献データベースの使用頻度が中国よりも高かった。しかし、HIS を利用した患者情報収集では、両国間に有意差は無かった。両国において、臨床研究支援のためのコンピュータの使用は十分ではなく、臨床医は臨床研究支援のために、HIS 上での包括的な診療録や HIS 上から全患者情報の使用を望んでいる

ことが明らかになった。

第5章では、日本における根拠に基づく医療 (Evidence Based Medicine: EBM) に対する臨床医の意識と問題点、EBM を実践するために必要な外来診療時の臨床医の情報ニーズを明らかにするために、北海道大学病院に勤務する 100 名の臨床医を対象とした調査結果をまとめた。質問の内容は、臨床問題の種類と発生頻度、問題解決のための情報源の種類と使用頻度、EBM を実践するための HIS に期待する情報支援機能、とした。その結果、日本の多くの臨床医は EBM の実践を試みているが、EBM のための情報学的手法が十分に普及していないことが明らかになった。

公开发表では、聴衆者約 20 名の参加を得て第 2 章を中心として約 30 分間の発表が行われた。これに対し副査の北島顕教授、寺沢浩一教授、主査の桜井恒太郎より本システムの評価、臨床応用、今後の進展に関する質問があり、以下の 7 点について約 25 分間の質疑応答があった。

本システムの評価に関しては、(1)開発したシステムの評価に関して、どのように医学概念の組み合わせの関係の強さについて評価したのか、(2)専門臨床医による用語の妥当性の判定の他に評価を行っていないのか、との質問があった。(1)の質問に対して、本システムの医学用語の形式概念分析は自動化されており主観的な概念が入りにくいこと、およびクロス表を用いて概念を正確かつ強力に記述していることを回答した。(2)では、情報学的な評価手法の可能性について言及するとともに、臨床情報の質の観点から臨床医による評価が必要なことを回答した。

臨床応用に関しては、(3)オントロジー支援システムは、患者の治療と結果(アウトカム)について知識の表現について使用できるか、(4)結果(アウトカム)若しくはそれらに関連する医学概念の抽出は行ったのか、(5)研究で用いた医学辞書が治療手順まで利用可能であり、結果(アウトカム)の知識がこのシステムで表現できたと仮定した際、本研究の結果に影響を与えるか、との質問があった。これらの質問に対して、(3)本システムはカルテなどから治療および結果(アウトカム)についての的確に概念を抽出できれば知識表現が可能なこと、(4)本研究の自然言語処理に使用した医学辞書(MEDIS version 2.0)が、診断用語および解剖用語のみであり、治療に関する概念を含んでいないため、医学用語を混合するアルゴリズムを構築したこと、(5)治療手順まで利用可能な医学辞書であれば、更に正確な治療や結果(outcome)に関する知識表現が可能であることを、プレゼンテーションの図を用いながら的確に回答した。

本研究の今後の進展に関する質問に関しては、(6)本博士論文では、オントロジー構築支援システムから臨床医が必要としている病院情報システムまで検討した中で、臨床医が EBM を実践するためにはどのような臨床情報システムが必要か、(7)本研究を中国の医療情報の中にどう役に立てたいか、との質問があった。(6)の質問に関しては、情報学的の視点から、EBM の実践のための患者集団からの根拠の生成のためには全国的・世界的なデータベースの構築が必要となると考えている、と将来展望が回答された。また、(7)に関しては、中国での近年の急速に医療の情報化が進んでいるが、それと共に医療情報倫理や医療の情報化が医療コストや医療の質にもたらす影響など日本と同様な問題が出現されると予想される。中国においても、これらの問題に対応できるような医療情報学分野の研究者、指導者が必要であろう、との回答があった。

本論文は、臨床医が EBM を実践するために必要とする情報ニーズを明らかにするとともに、日本語での医学オントロジーによる臨床知識の表現と処理を通じて、今後の EBM 支援を可能とする病院情報システムを検討したしたものとして高く評価される。

審査員一同は、これらの成果を高く評価し、大学院課程における研鑽や取得単位なども併せ申請者が博士 (医学) の学位を受けるのに十分な資格を有するものと判定した。