

## 「診断困難例ケースサーチ：J-CaseMap」の導入にあたって

内科学は多くの疾患を対象とするため、総合的に学ぶことは大変困難です。このため内科学の専門分化が進んできましたが、一方で総合医の養成を疎かにすることはできません。内科学を総合的に学ぶ一つの方法が、多彩な症例の学習であることは申すまでもありません。

日本内科学会地方会は、多くの症例を学ぶことを目的として戦後に活動をはじめました。例えば、関東地方会は1946年（昭和21年）に組織され、現在までに約660回、毎年9回の開催により年間約1000例の症例を報告してきました。1990年代以後、全国の地方会症例報告はデジタル化され、総数は6万例を超えています。しかしながら当初の検索システムは機能が高くなかったために、私が日本内科学会理事長を務めていた2009年、東京大学医学部附属病院医療情報部の大江和彦教授、荒牧英治助教（現奈良先端科学技術大学院大学准教授）に依頼し、高速検索システムである「症例くん」を作成していただきました。これは症例検索に大きな力を発揮し、現在でも毎月千件以上のアクセスがあります。

しかしながら「症例くん」は、症例報告に記載された医学用語の検索システムであり、陽性所見と陰性所見を区別したり、症例の文脈を読んだりすることができません。一方、現在、自然言語の機械学習が注目されていますが、今の深層学習では、症例報告の文脈を理解することはまだできません。そこで2016年に、AMED臨床研究等ICT基盤構築研究事業「人工知能による総合診療診断支援システムの開発」（研究代表者：永井良三）の支援を受け、症例報告の文脈と鍵となる医学用語の関係をグラフ表示し、これを検索や診断支援のためのシステムとしたのが「診断困難例ケースサーチ：J-CaseMap」です。また医学用語を機械が理解するには、少なくとも数万語の用語リストと、それぞれの用語の同義語、類義語、検索のための用語の辞書を用意する必要がありますが、あわせてその作業も行いました。

まず自治医科大学の医師をはじめとする約150名の内科医が中心となり、約2万例の症例の論理を図式化しました。しかし必ずしも統一が取れていなかったため、私自身が修正した6000例を利用してシステムを構築しました。検索ソフトは小田啓太（元Googleエンジニア、自治医科大学）、今井健（東京大学、自治医科大学）、佐藤寿彦（プレジジョン社、医師、本学会員）が分担して開発しました。検索アルゴリズムの一部は、日本内科学会と開発者が共同出願し特許登録されました（特許6539818、）。その後の開発には、AMED事業以外に、令和元年度NEDO Connected Industries 推進のための協調領域デー

タ共有・AI システム開発促進事業「医療情報を横断的に統合した診療支援 AI システムの開発」(研究代表者:佐藤寿彦、分担研究者:永井良三)の支援を受けております。

大量の症例報告を構造化すると、未診断症例の類似症例の検索や、病態の組み合わせによる原因疾患(病態)の推測が可能となります。しかしこれは確定的診断を行うものではなく、鑑別診断の参考となる疾患や病態をリマインドするためのシステムと位置づけられます。

内科学の情報統合も学会が担う重要な使命です。概念や用語の関係性(因果、包含関係、従属関係など)を体系化するのが、オントロジー(分類体系)です。古典的な学術ですが、21世紀のオントロジーは、医学知識の構造化と次世代AIの基盤として必須です。そのためすでに世界中でいくつかの重要な取り組みが行われています。世界標準とされる疾病分類「疾病及び関連保健問題の国際統計分類(国際疾病分類ICD)」もその一つです。しかしこれは国際的に統一した基準で定められた死因及び疾病の分類であり、きめ細かい臨床現場の記述には限界があります。世界最大の医学用語集は、30か国以上が参加して進めるSNOMED CT(Systematized Nomenclature of Medicine Clinical Terms)です。これは国際機関により管理されており、80万の用語で35万件以上の医学概念を記述しています。同義語、階層構造、意味リンクによって概念同士が関連付けられ、ICD-11とも連携する予定です。しかし膨大なコスト負担を求められるため日本は参加しておりません。今後、医療データを活用するには、単なるSNOMEDの日本語翻訳ではなく、日本独自の医学用語集を編纂する必要があります。

わが国では大江教授らが、「医療情報システムのための医療情報知識基盤研究開発事業」を進めてきました。この研究では疾患概念の定義から始め、そのもとに病態、症状、身体所見、検査所見が因果の連鎖として記述されます。また解剖構造も情報として付加され、疾患概念は連鎖図の局所的な面として表現されます。これは合理的に考えられた概念を先行させるトップダウン型のアプローチです。一方、今回導入する内科学会のシステムは、症例報告の文脈を辿りつつ病態と徴候を階層化し、用語間の関連図を作成する経験主義的なボトムアップ型アプローチをとっています。症例報告は、病態、症状、所見に関する医学用語が豊富です。これと大江教授らのトップダウンの医学用語体系を統合し、ICD-11と連結した時に、はじめて機械が学習できる日本語の医学用語体系が構築されます。また症例をもとに医学用語の概念を構築する活動は、内科学の全体性回復につながると考えております。

まだ6000例の症例報告を用いた試行ですが、日本内科学会の会員の皆様に活用いただければ幸いです。