臨床医の診断推論能力の客観的評価と国際比較

藤田保健衛生大学一般内科 助教授

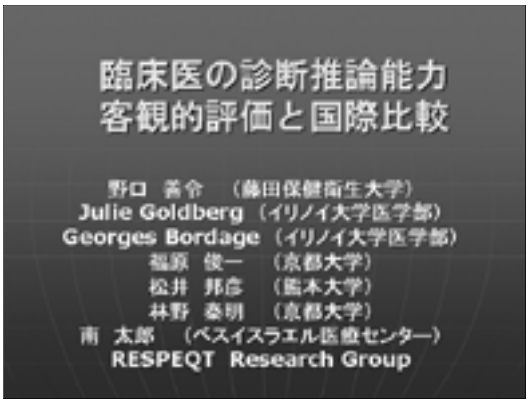
野口　善令

【スライド-1, 2】

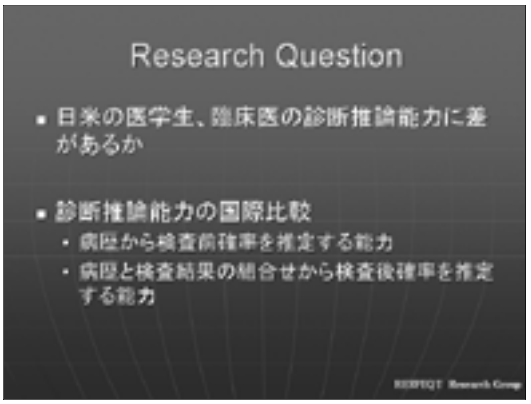
北米は臨床教育の先進地域だということになっておりまして、北米の教育方法によれば臨床能力の高い医者を効率的に育て上げることができると言われております。そこで今回我々は、日米の医学生、臨床医の診断推論能力に差があるかどうかということを国際比較しました。

診断推論能力の要素の内で、病歴から検査前確率を推定する能力、それから、病歴と検査結果の組み合せから検査後確率を推定する能力を測定して、国際比較しました。

スライド-1



スライド-2

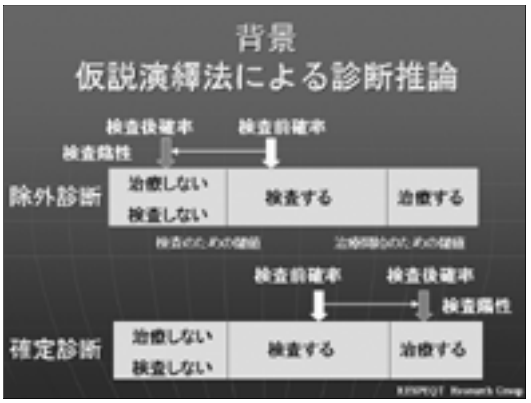


【スライド-3】

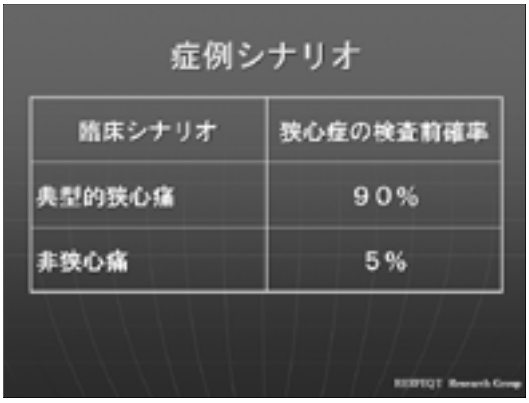
背景として、仮説演繹法による診断推論について少し説明します。

仮説演繹法による診断推論では、まず、病歴から患者さんがある疾患を持っている可能性がどれ位かという検査前確率を推定します。次に、その検査前確率に検査結果を加えて、検査後確率を推定します。例えば患者さんの検査結果が陰性であるとすると、検査前確率・・・つまり患者さんがある疾患を持つ確率が下がって、検査後確率が左向きの矢印の所に来ます。それで、これ以上の検索を放棄してよいというレベルまで達すれば、除外診断ということになります。逆に検査結果が陽性と出て、検査後確率が上がって、治療してよいという判断できるレベルになれば確定診断となります。今回の研究では、この考え方の枠組みに基づいて、臨床医の診断推論能力を評価しました。

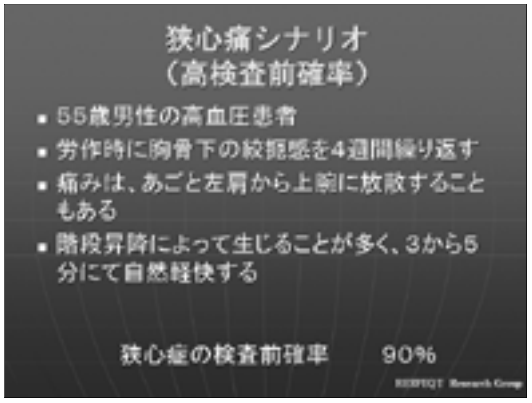
スライド-3



スライド-4



スライド-5



【スライド-4】

実際の方法ですが、症例シナリオを提示して、その後で、検査前確率、検査後確率を回答してもらうという方法を取りました。

まず症例シナリオですが、狭心痛の検査前確率が90％のものを典型的狭心痛シナリオ、それから、狭心症の検査前確率5%のものを非狭心痛シナリオというように、2種類用意しました。

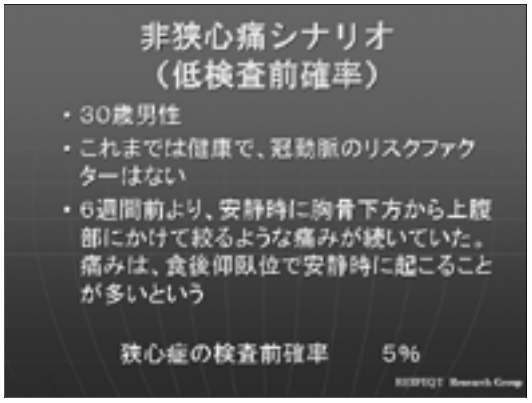
【スライド-5】

これは具体的な狭心痛シナリオで、高い検査前確率のシナリオの内容です。55歳男性の高血圧患者。労作時に胸骨下の絞扼感を4週間繰り返す。痛みは、あごと左肩から上腕に放散することもある。階段昇降によって生じることが多く、3分から5分で自然軽快する。こういうような内容の症例を呈示しております。これは狭心症の検査前確率は90%程度であろうと推定されます。

【スライド-6】

こちらは非狭心痛シナリオで、低い方のシナリオの例です。30歳男性。これまでは健康で、冠動脈のリスクファクターはないと。6週間前から、安静時に胸骨下方から上腹部にかけて絞るような痛みが続いていた。痛みは、食後仰臥位で安静時に起こることが多いという。こういうような内容のシナリオです。これは、狭心症の検査前確率が5%程度であろうと推定されます。

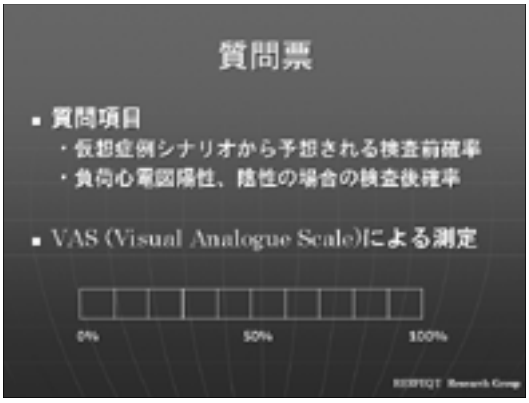
スライド-6



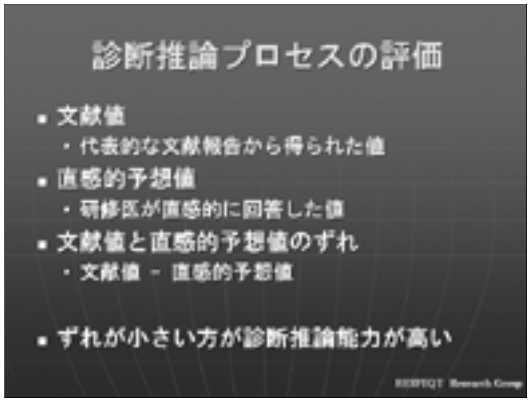
【スライド-7】

このような臨床シナリオを提示した後で、質問票を用いて、予想される検査前確率、それから負荷心電図が陽性の場合・陰性の場合の、それぞれの検査後確率を回答してもらいました。回答にはVAS（Visual Analogue Scale）を用い、このバーグラフの中で好きなところにチェックを入れてもらって回答してもらう

スライド-7



スライド-8



という方式を採りました。

【スライド-8】

診断推論プロセスの評価です。

文献値というのは、先ほどの症例シナリオで代表的な文献報告から推定される値で、これを文献値としました。それから直感的予想値は、研修医、医学生が直感的に回答した値です。ここで、文献値と直感的予想値のずれを計測して評価しております。

文献値というのは最良の診断推論から得られる予想というものを表わしており、このずれが小さい方が診断推論能力が高いと判定されます。

【スライド-9】

対象ですが、医学生に関しては、日本の3つの国公立大学医学部学生224人、それから米国に関してはイリノイ大学医学部シカゴ校の学生112名を対象に調査しました。

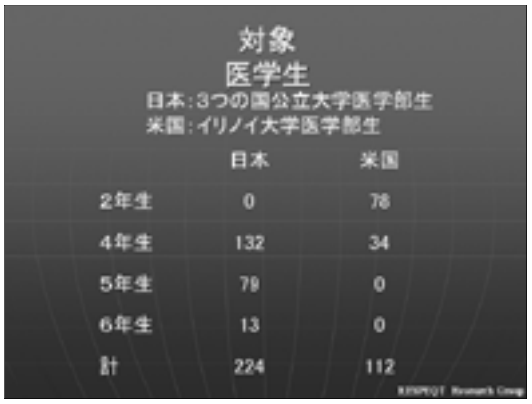
【スライド-10】

更に臨床医に関しては、日本だけですが、日本全国の11の教育指定病院、国公立大学、私立大学、国公立病院、民間病院など、色々含まれておりますけれども、総計で262名に調査が行なわれました。

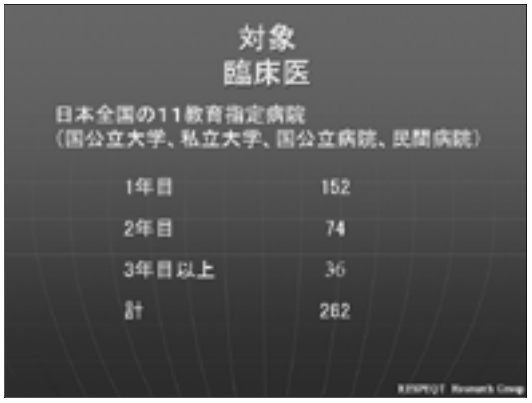
【スライド-11】

これは、まず、典型的狭心痛シナリオ、つまり検査前確率の高いシナリオの結果の国際比較です。左から、文献から得られた値、アメリカの学生、日本の学生と並んでおります。真ん中が検査前確率で、

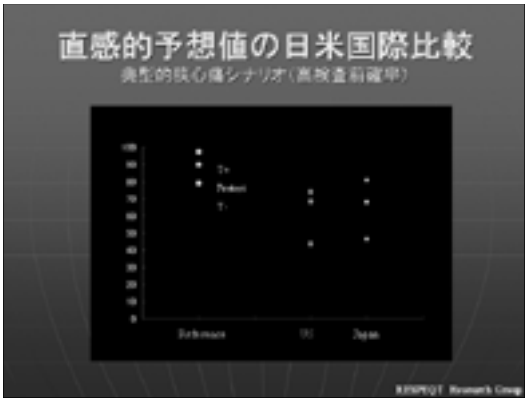
スライド-9



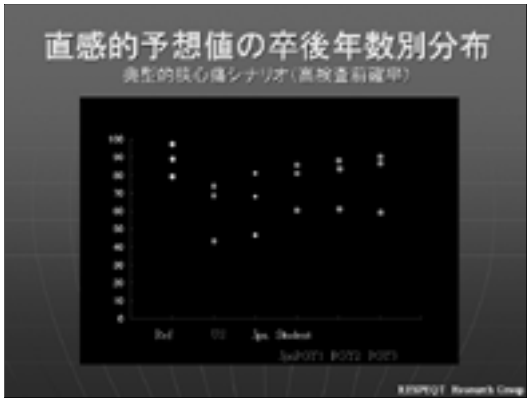
スライド-10



スライド-11



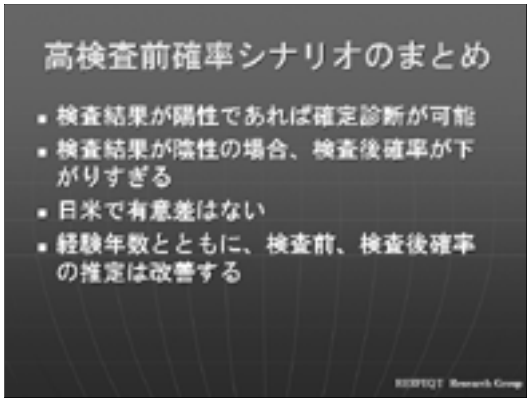
スライド-12



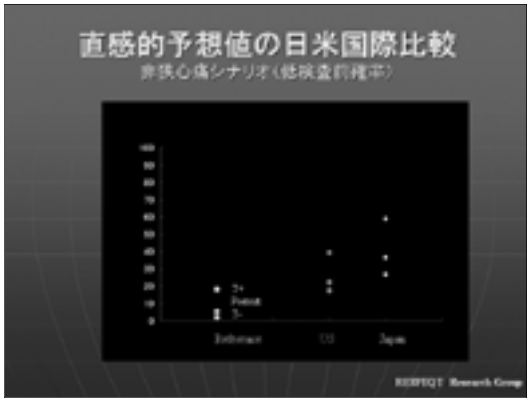
上が検査結果（負荷心電図）が陽性の場合の検査後の確率。それから、下が負荷心電図が陰性の場合の検査後確率というようにグラフで表わしてあります。

結果を見ますと、文献値より少し低い傾向はありますが、比較的良い推定をしている。日本とアメリカの間に少し差があるようですが、有意差はありませんでした。それから、検査結果が陰性と出ると、検査後確率を低すぎる方向に推定するという傾向が見られました。

スライド-13



スライド-14



【スライド-12】

これは、今挙げた日米の医学生の比較にプラスして、日本の研修医の卒後1年目、2年目、3年目を表わしたグラフです。傾向としては同じようです。卒後年次が上がるに連れて、だんだん、この推定は良くなっていくという方向に向かっています。

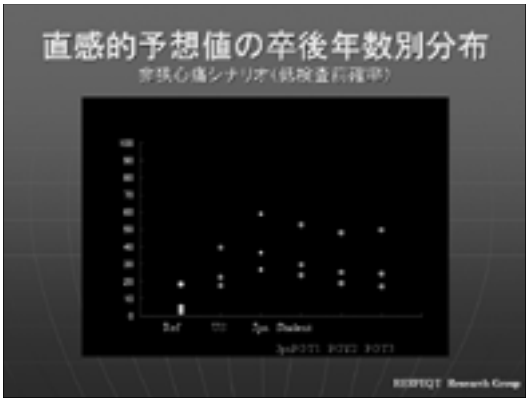
【スライド-13】

そこで高い検査前確率シナリオのまとめですが、検査結果が陽性であれば、確定診断は可能であろうと考えられます。検査結果が陰性の場合、検査後確率が下がりすぎる傾向があります。ただし、日米で有意差はありませんでした。ただ、卒後の経験年数がだんだん上がるとともに、検査前確率、検査後確率の推定は改善する傾向が見られました。

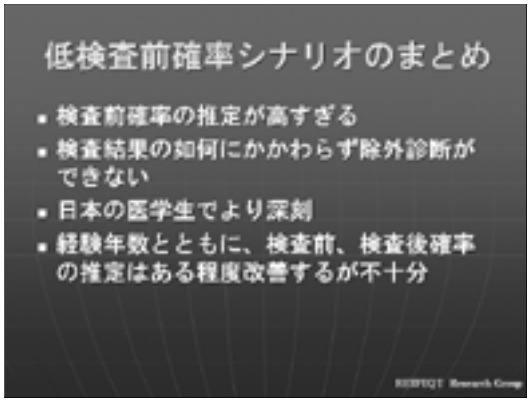
【スライド-14】

スライドは非狭心痛シナリオ、つまり検査前確率が低いシナリオの、日米の医学生

スライド-15



スライド-16



の国際比較です。左から文献値、アメリカ、日本の順に並んでいます。これを見ますと、文献から得られる値に比べて、検査前確率、検査後確率ともに高い方向に推定しております。この場合は、日本とアメリカの間で有意差がありました。日本の学生の方が、より外れた方向で推定しておりました。

【スライド-15】

これも同じように日本の研修医、卒後1年目、2年目、3年目以上というデーターを付け加えてグラフにしたものですが、この場合も、卒後年数が上がると文献値に近づく。少しずつ改善していく傾向は認めますが、改善の程度は充分ではないと判断されました。

【スライド-16】

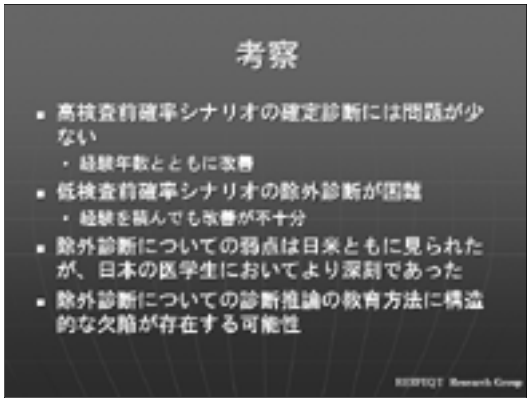
低い検査前確率シナリオのまとめですが、検査前確率の推定が高すぎる傾向があります。そのために、検査結果の如何に関わらず、除外診断ができないという問題があります。この問題は、日本の医学生でより深刻でした。卒後年数とともに検査前確率、検査後確率ともに推定はある程度改善しますが、不充分であると考えられています。

【スライド-17】

全体の考察ですが、高い検査前確率のシナリオの確定診断には、比較的問題が少ないと考えられました。更にこれは経験年数とともに改善いたします。低い検査前確率シナリオの除外診断は困難であろうと推定します。これは問題として、経験を積んでも改善が不充分であるということが認められました。この除外診断についての弱点は日米ともに見られましたが、医学生で比較すると、日本の医学生でより深刻でした。

全体の結論として、除外診断についての診断推論の教育方法に構造的な欠陥が存在する可能性があります。この除外診

スライド-17



断の弱点というのは、患者さんを診た場合に、除外診断ができなくていたずらに検査を繰り返すという診療行動を、うまく説明できるモデルではないかと考えております。更にアメリカの学生の方が、弱点がある日本医学生よりも、良い推定をしているということで、アメリカの卒後教育並びに卒前教育の教育方法から、何か日本が改善をする点で学ぶことがあるのではないかという結論を持っております。

質疑応答

会場： 興味深い発表、ありがとうございました。特に後半のところの、除外診断ができる思考プロセスないしは類推するプロセスが充分身に付いていないために、実は無駄に医療費を重ねているのではないかという指摘は、非常に的を射ており、私も同感なのですが、教育プログラムの中で、鑑別診断とか、あるいは病体生理の教育のあたりのところについて、もし現時点で先生のお考え等ございましたら、ご披露頂ければありがたいのですが。

野口： まず鑑別診断で、検査前確率を考え、鑑別診断をまずきちんと上げて、その鑑別診断の一つひとつについて検査前確率を考えて、更に検査をした場合に、検査後確率を考えるというような思考の枠組みが、果たして教育されているかどうかという問題があるかと思います。これは定量的に調べたわけではないのですが、アメリカの教育と比べて一番強く日本の教育に足りないと感じる点が、やはりこの鑑別診断をするという考え方の枠組みをもって、それに基づいてトレーニングしているかどうかということではないかなという印象を持っております。

会場： このスタディで、検査前の高い確率とかいう症例は何例位ずつお使いになっているのですか。

野口： シナリオは今挙げた2つと、もう一つ、中間のシナリオというのを作りましたので、3つです。

そのデーターをどこから取っているかというと、アメリカにこういう症状があった場合に、どれ位の狭心症の確率なのかを推定したようなスタディがもうずいぶん出ておりましたので、そういうものからデーターを取って、ああいうシナリオを作りました。

会場： そうなんですか。正直な話、狭心症はもの凄く訴えが多彩です。いろいろな症例があり、その症例毎の診断確率というものがやはりあるのです。先生が90％で典型的狭心症となさっても、その間にも色々な症例があるのです。ですから、本当だったら、もう少しそういうところを踏まえた詳しいデーターが欲

しいなと、ちょっと思ったのです。特に私は循環器を専門にしておりますものですから、診断する時にはそんなに簡単ではないという感じがいたしました。

野口： 確かに現場ではおっしゃる通りですが、ある程度モデル化しないと、こういう定量的な評価ができませんので、あえて単純化して、非常に高いだろうという症例と非常に低いだろうと症例のシナリオについてやってみたというところです。

会場： 先ほどのご質問と多少関係があるかもしれませんが、背景として、日米の狭心症の罹患率、有病率とかが、この研究に何かバイアスを与えているかどうかということが一つです。あと一つは、狭心症についてのシナリオですが、他の、例えば私のやっている消化器系をやった場合はどうか。そういった点はどうなのでしょうか。

野口： まず一番目のご質問ですが、確かに日米で罹患率、有病率に差があって、日本だと少し低い方向にずれるのではないかと思います。もしもその経験を反映して、学生とか研修医が忠実に推定していれば、どちらかと言うと、日本の方がより低い方向に推定してもよいものが、逆の方向に行っているということで、そこら辺から、有病率によるバイアスというのはあまりないのではないかと考えております。

それから二番目の、他の領域の疾患ではどうかということですが、実は他の領域では、どんな症状があったら疾患の検査前確率がどれ位かというデーターがあまり無いのです。非常に興味があることで、例えば胃潰瘍だったらどうかというようなこともやってみたいのですが、どうもうまくシナリオができないので、それに関しては手を付けておりません。

会場： 先ほどの会場からの質問にも関係あるのですが、これはあまりにも問題を単純化し過ぎていると思うのです。それで、私は、これだけでもっていろいろなことを言われるのには、すごく危険があるという気がするのです。こういうことについて、これだけのデーターを集められたということについては敬意を払うのですが、ただ、あまりにもこういう単純化したモデルで、診断能力であるとか、日米の臨床医の能力の差だとかという辺りまで、どんどん話を広げて議論するということは、むしろやめるべきであるというように思いますので、その点だけはちょっとご注意をいただいた方がよいのではないかなと思いました。

座長： 日本が国際化していく中で、こういう問題が、アメリカの学生と日本の学生の共通基盤として認識される時代が来ると思うのです。その意味でこういった研究は非常に重要だと思うのですけれども、お聞きしたいのは、先生がこういう研究をされようと思ったそもそもの動機は何なのでしょうか。

野口： 動機は、先ほども言いましたように、鑑別診断ということの考え方の枠組みを持って教育しているかどうか、そしてその結果が、実際どういうふうに影響しているのかなという疑問を持ったことでした。

座長： 言葉を換えると、日本の医学生にはその枠組みが無いと。

野口： 無いと言うと、先ほどのご指摘のような、あまりに単純化・一般化された言い方になってしまいますが、日本とアメリカの両者を見ていると、やはり足りないのではないかなという印象は受けます。