患者向け医療情報の信頼性の評価に関する国際研究
－DISCERN 日本語版の開発とアトピー情報の評価　および国際比較－

国立成育医療センター第一専門診療部アレルギー科　医長
大矢　幸弘

【スライド-1】

始めに、このような研究と発表の機会を与えてくださいましたファイザーヘルスリサーチ振興財団の皆様方に感謝申し上げます。

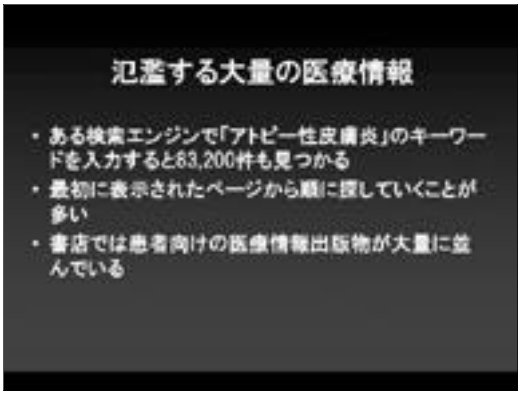
スライド-1

【スライド-2】

現在は、氾濫する大量の医療情報という世界に我々は生きています。

ある検索エンジンで「アトピー性皮膚炎」とキーワードを入れます。そうすると8万3,200件も引っ掛かってきます。患者さん達は少しでも良い情報を求めて探しているわけですけれども、全部読めるわけはないし、最初に表示されたページから順に探していくことが多くなるわけです。それから、書店では患者向けの医療情報の出版物が大量に並んでいます。私の専門分野でありますアレルギー疾患、特にアトピー性皮膚炎などは、ものすごい数が並んでいまして、どれを選んでよいのかわからないというような状況がございます。

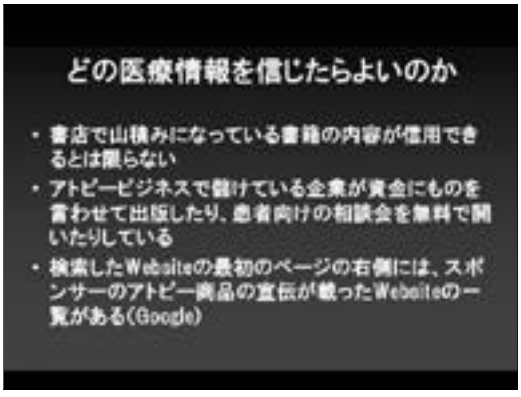
スライド-2



【スライド-3】

また、書店で山積みになっている本があります。そういう本が内容が良いのかと言いますと、必ずしもそうではありません。アトピービジネスで儲けている企業が資金にものを言わせて出版したり、患者向けの相談会を無料で開いたりしております。ですから横積みになっている本というのは、意外と危険なことがある

スライド-3



わけです。でも、それはなかなかわかりませんから、結構引っ掛かってしまうわけです。

またウェブサイトも、日本のウェブサイトは何故か最初のページあるいは次のページの右側の方にはスポンサーのアトピー商品の宣伝が載っているのです。ということで、最初のページを開くと宣伝がわーっと出てきますが、そちらを開いてしまうと大変なことになるわけです。

【スライド-4】

研究の背景としては、こうした氾濫した医療情報があって、また、医師によって言うことが違うという問題があります。どの医療情報を信じたらよいのかわからないということがありますので、

スライド-4

【スライド-5】

では、患者向けの医療情報の信頼性を専門医以外の日本の人々がチェックできるツールを開発する必要があるのではないかと考えました。

そういうものは世界には他に無いのかということで探しましたら、オックスフォード大学で作成されたDISCERNというものがございました。そこで、DISCERN instrumentの日本語版を開発し、その性能を調べてみようということで、この国際共同研究を始めたわけです。

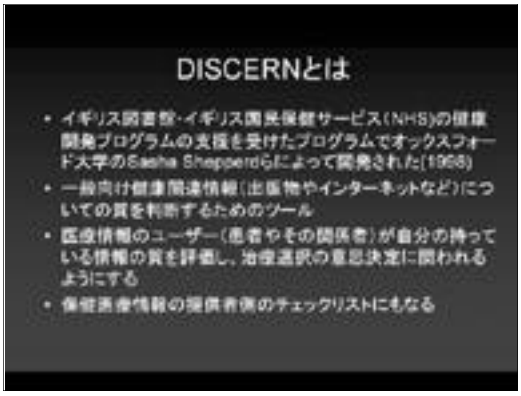
日英のアトピー性皮膚炎の情報をDISCERNを使って調べてみたらどうなるかということもやってみました。

スライド-5

【スライド-6】

このDISCERNというものですが、これはイギリス図書館、イギリス国民保健サービス（NHS）の健康開発プログラムの支援を受けたプログラムで、オックスフォード大学のSasha Shepperdらによって開発されています。それが1998年のことです。

スライド-6



一般向けの健康関連情報、出版物とかインターネットなどについての質を判断するためのツールです。医療情報のユーザー、つまり患者とか関係者が自分の持っている

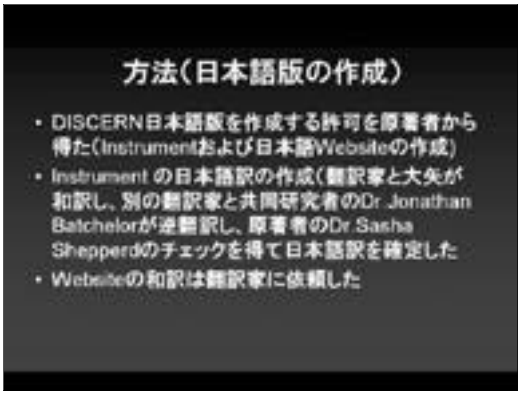
情報の質を評価し、治療選択の意思決定に関われるようにするという目的がございます。そして、これは保健医療情報の提供者側のチェックリストにもなります。ですから、このDISCERNを見ることによって、少しでも質の高い医療情報を提供するような作り方をすることもできるわけです。

スライド-7

【スライド-7】

DISCERNのウェブサイトを開きますと、こういう画面が出てまいります。これをそのままコピーして日本語化したものを私どもは作りました。日本語版はまだ公開されていませんが、いずれ専門のウェブサイトを立ち上げたいと考えております。

スライド-8

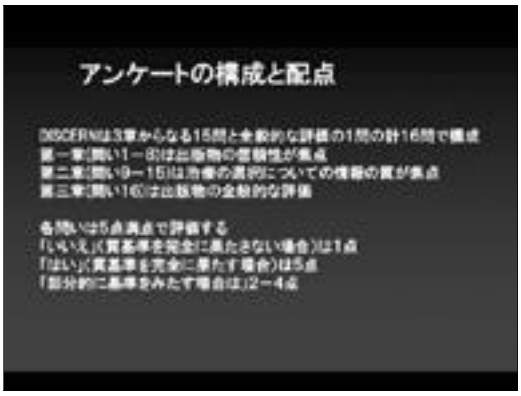


【スライド-8】

日本語版作成をどのようにやったかということですが、まず原著者からこのInstrumentの日本語版及び日本語のウェブサイトの作成をすることの許可を得ました。

そしてInstrumentの日本語訳の作成をしました。翻訳家と私が和訳して、またそれを何人かの患者さんを含めたいろいろな人に見てもらって、だいたい理解できる内容かということをもう一回チェックして、そして、別の翻訳家とそれから共同研究者のイギリスの医師のJonathan Batchelorがそれを逆翻訳し、逆翻訳したものを原著者のDr. Sasha Shepperdにチェックしてもらいました。このプロセスを繰り返して、一応イギリスで出しているものと日本で出しているものと、だいたい同じ意味であろうという、いわゆる翻訳妥当性を検証いたしました。

スライド-9



【スライド-9】

Instrumentというのは、いわゆるアンケートによってチェックしていく採点表なのですけれども、これは三章からなっ

Instrumentというのが採点するツールなのですが、これ以外にウェブサイトでいろいろな解説がございます。そういうものはプロの翻訳家に依頼してやってもらいました。

ています。15問と最後に全般的な評価をする1問で、計16問から構成されています。

第一章は、出版物の信頼性に関するものが焦点。

それから、第二章は治療の選択についての情報の質が焦点になっています。

第三章は全般的な評価ということで、5点満点です。質が明らかに悪く、評価が基準を満たしていないと思われるものには1点、完全に基準を満たしているものには5点、部分的な場合には2点から4点ということで点数を付けていきます。

【スライド-10】

スライド-10

例えば第4問などは、このようになっています。

「このパンフレットを編集するのに使われた著者や製作者以外の情報源は明らかか。」そこでどのように採点するかというと、ヒントがあります。

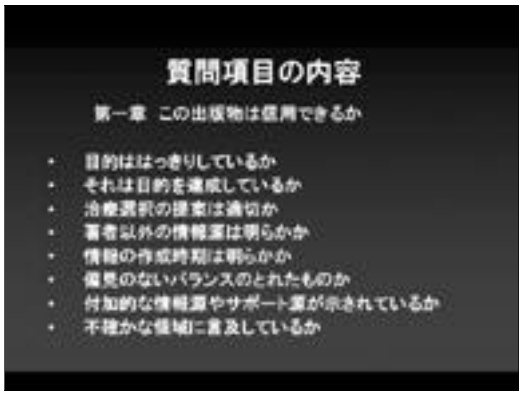
「この治療選択に関して書かれた主張や叙述には、根拠として使用されている情報源の言及を伴っているかどうか」ということです。勝手に主観的に書いているのではないかどうか、ちゃんとした引用文献を、どこから取って何を根拠にそういう情報を載せているのかということが重要になってくるわけです。まさに参考文献というものの質を評価するわけです。いわゆるEBMの基準の高い引用文献かどうかということがこれで問題になってくるのですが、そんなことでチェックしていきますと、まともな情報源からこれが得られているのかどうかということがだいたいわかってくるわけです。

「5点満点のスコアを得るためには、両方のヒントを満たさなくてはいけない。付加的な情報源やサポート源のリストはこの出版物の根拠として必要な情報源ではない。」とされています。

問7にもちょっと似た質問があります。付加的な情報源とかサポート源のリストについても、チェックをしていくということになっています。

【スライド-11】

スライド-11



質問の項目の内容ですが、第一章はこの出版物が信用できるかどうかということになっています。

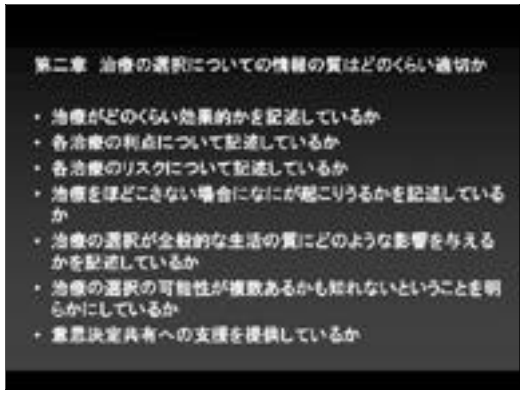
第一問から八問までございまして、

- 情報の目的ははっきりしているか
- その目的をパンフレットなりウェブサイトは達成しているかどうか
- そして治療選択の提案は適切か

どうか
・著者以外の情報源は明らかか
・情報の作成時期は明らかであるか
・偏見のないバランスのとれたものであるか
・付加的な情報源やサポート源が示されているか
・不確かな領域に言及しているか
ということです。

【スライド-12】

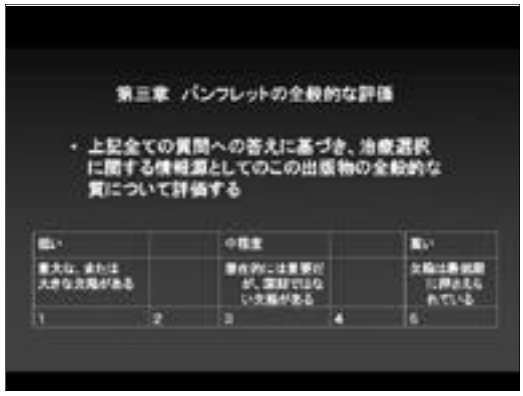
スライド-12



そして第二章になりますと、治療の選択についての情報の質はどれくらい適切かということで、
・治療がどのくらい効果的かを記述しているか
・各治療の利点について記述しているか
・各治療のリスクについて記述しているか
・治療を施さない場合に何が起こりうるかを記述しているか
・治療の選択が全般的な生活の質にどのような影響を与えるかを記述しているか
・治療の選択の可能性が複数あるかもしれないということを明らかにしているか
・意思決定共有への支援を提供しているか
「なかなか治療を施さない場合に何が起こるかを記述しているか」というあたりは、あまり書いていないウェブサイトあるいはパンフレットが多いように思いました。

【スライド-13】

スライド-13



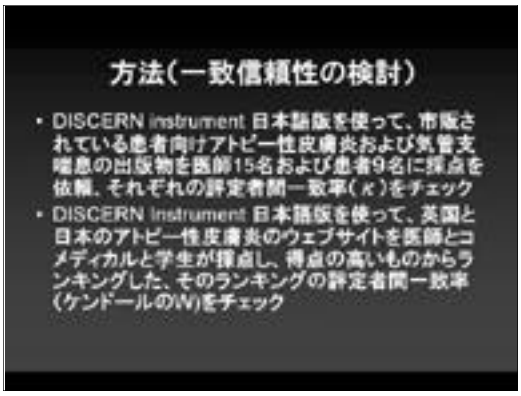
最後に、全ての質問への答えに基づき、治療選択に関する情報源としてのこの出版物の全般的な質について評価致します。
重大な欠陥があるものは1点で、欠陥は最小限に抑えられているものは5点という具合につけていくわけです。

【スライド-14】

そこで、どのくらいこれが採点者による一致信頼性があるのかということを調べてみました。
市販されている患者向けのアトピー性皮膚炎と気管支喘息の出版物を、医師15名および患者9名に採点をしてもらい、それぞれの評定者間の一致率をチェックいたしました。
これが結構時間がかかります。なにしろその本を1冊読まなければいけないし、ウ

ェブサイトを見なければいけないので、実はこれを頼むと、皆すごくいやがりました。頼んでから3カ月経っても出さない人がいまして、15名揃うのに結局、半年ぐらいかかりました。そんなことで、随分時間がかかってしまったのですが、その一致率をチェックしていきました。

スライド-14



それから、このInstrumentを使いまして、イギリスと日本のアトピー性皮膚炎のウェブサイトを採点しました。高いものから順番にランキングしたものを後でお見せしようと思います。そして、評定者間の一致率も見てみました。

【スライド-15】

アトピー性皮膚炎で使ったパンフレットはこれです。確か60ページぐらいあると思います。

【スライド-16】

こちらが環境省の外郭団体の公健協会が作っているもので、まともなパンフレットですが、これも30数ページあります。

スライド-15

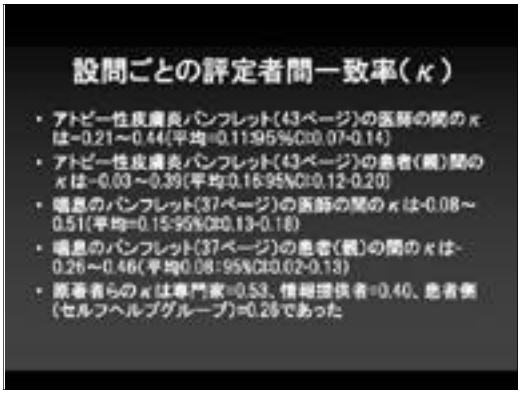
スライド-16

【スライド-17】

それで、どのくらい評定者間の一致率があるか。

一致率はκ（カッパー）というものを用いています。これは、ある人が5点とつけて、ある人が3点とつけたら全然一致していないということになります。ある人が5点ならこっちも5点とつけなければいけない。ですから、16問全部にこの

スライド-17



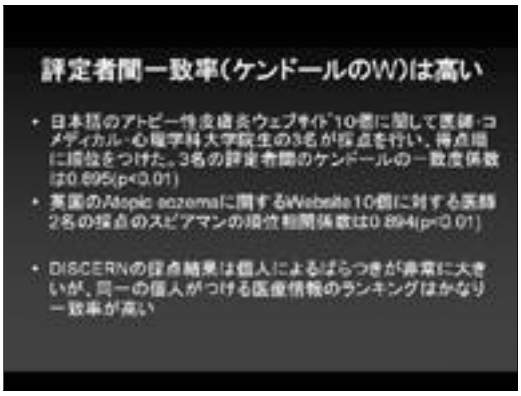
パンフレットに5点付き、別の人も5点付いたら満点です。そういう場合には1になります。それから、てんでばらばらで、それが偶然の確率よりも低い場合には－（マイナス）が付いてしまいます。偶然の確率で合っているぐらいの程度しかない場合は0になります。そういうような一致率なので、かなり低くなることが予想されたのですが、実際やはり低かったです。

アトピー性皮膚炎のパンフレットは、医師の間のκは－0.21から0.44ということでした。0.44まで高い組み合わせもあったので、合うときには合うのだと思いましたが、平均でいきますと0.11ということで、あまり高い一致率は得られませんでした。各設問についてある人は5点つけているのに、ある人は1点つけてあるから「どうなっているのだろう」というようなものも中にはありました。それぞれ医師、患者、それから喘息の患者の親御さんにそれぞれやってもらいましたが、あまり差がありませんでした。

ところが、原著者らの論文を見ますと、専門家が付けた場合は0.53の一致率で、これは高いです。それから情報提供者が0.4、患者側のセルフヘルプグループにやらせたものが0.2ということで、専門家がやれば高いという数字が出ています。しかし、我々の間は専門家がやっても全然高くありませんでした。それで、あまり信頼性がないのではないかと思って、実は困っていました。今のパンフレット以外にも、ものすごくたくさんやってもらいました。他のウェブサイトを見たり、パンフレットをいくつかやったり、何度も何度も試行錯誤したのですが、何回やりましてもなかなか一致率が高くなりませんでした。

【スライド-18】

スライド-18



もうこのツールは駄目なのかなと思いましたが、実はそうではありませんでした。日本語のアトピー性皮膚炎のウェブサイト10個に関して、医師とコメディカル、心理学科の大学院生の3人が採点を行いました。そして、得点順に順位をつけました。得点の高い物が信頼性が高いということになるわけですけれども、あるものは70点、低い物は10何点というようなことで、順番に高い物から順位をつけたわけです。その3名の評定者間のケンドールの一致率を見たら非常に高く、0.985だったのです。

イギリスのウェブサイトも医師二人で採点してみました。そうしたらこちらも、2名の場合はスピアマンの順位相関係数になり、ケンドールの一致率と同じことですが、これも0.894で非常に高い結果が得られました。

【スライド-19】

つまりどういうことかと言うと、DISCERNの採点結果というのは個人によるばらつきが非常に大きいのですが、同一の個人が付けた医療情報のランキングはかなり一致

率が高いということがわかったわけです。つまり、ある人は非常に辛い点をつける傾向はあるけれども、ある人は非常に甘い点を付ける傾向がある。だからそれぞれ違った人が付けた点数というのはあまり一致しないのだけれども、同じ人にいくつかの情報を付けさせてランキングをすると、だいたい一致する。そういうことがわかりました。ですから、κは低いけれども、ケンドールのWは高いという結果が出たわけです。

ですから、得点の絶対値そのものの信頼性は低い。そして原著者は訓練によって一致率のκを高くしています。でもそのような手続きは実用的ではありません。しかし同一の個人が採点した複数のウェブサイトのランキングに関しては、このように異なる評定者の間でも一致率が高いということで、別に訓練をしなくてもこのようにランキングをつけることはできるということです。

【スライド-20】

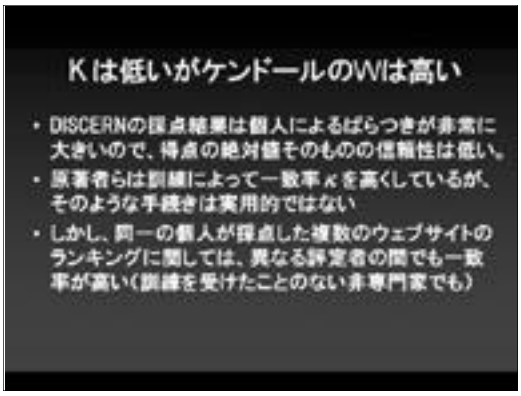
それから日英のアトピー性皮膚炎のウェブサイトを検索し、上位10個を比較してみました。

同じ人間がやらないと比べられないのですが、評定者のBはイギリスのウェブサイトと日本のウェブサイトの両方とも採点しています。ですからそこをちょっと見ていただきたいと思います。そうするとイギリスの方が高いという結果が出ました。ばらつきも少なかった。

【スライド-21】

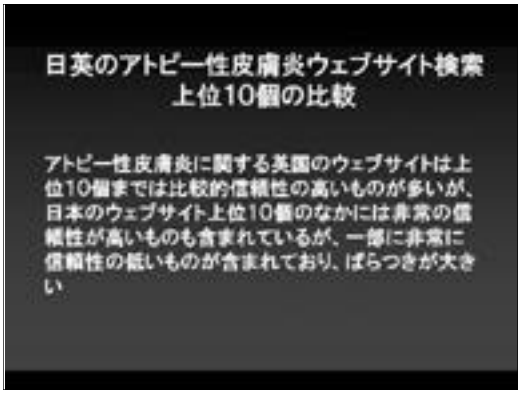
ということで、アトピー性皮膚炎に関する英国のウェブサイトは上位10個までは比較的信頼性の高いものが多いが、日本の場合は、信頼性の非常に高い物もあるけれども低い物も含まれており、ばら

スライド-19



スライド-20

スライド-21



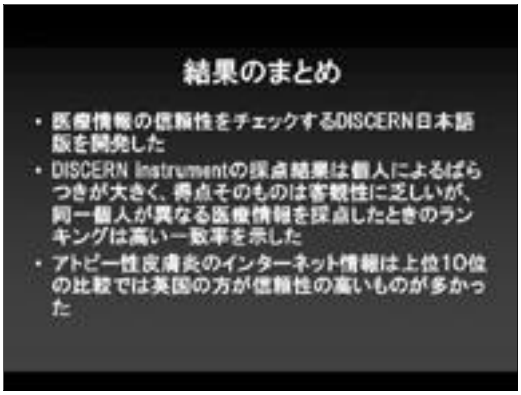
スライド-22

つきが大きかったです。

【スライド-22】
　これは日本のウェブサイトで最も高かったものを示しています。

【スライド-23】
　結果のまとめです。
　DISCERNの日本版を開発しました。
　先ほどのように、得点そのものの客観性は乏しいですけれども、ランキングに関しては結構信頼できるということです。

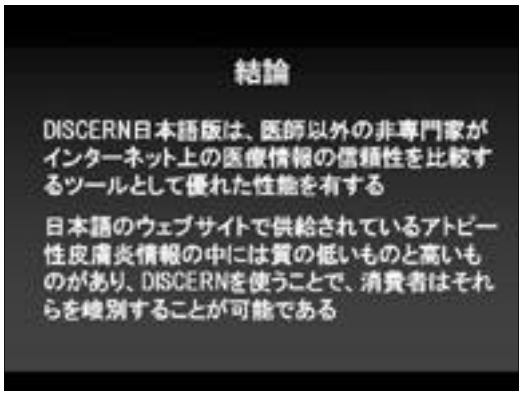
スライド-23



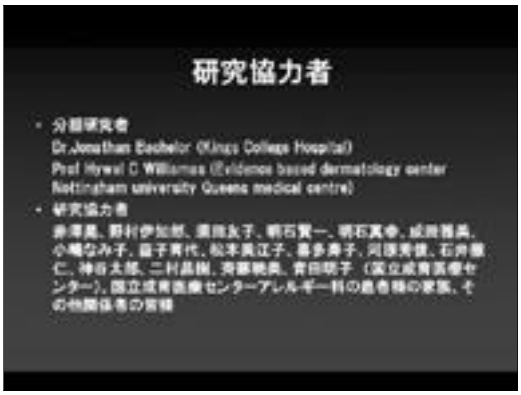
【スライド-24】
　結論です。インターネット上の医療情報の信頼性の絶対性のツールではなく、信頼性を比較するツールとしては優れた性能を有していると思います。
　日本のウェブサイトで供給されている情報の中には質の低いものも高いものもありますので、このDISCERNを使うことで、消費者はそれらを峻別することが可能であると思います。

【スライド-25】
　研究協力者です。上の2名が分担研究者であります。
　たくさん方々にご協力をいただきました。

スライド-24



スライド-25



質疑応答

会場：　医療機関の情報開示のツールとしてホームページに治療成績を出している他、私どもは、非常に内容の良いページは、インフォームドコンセントの場でも活用した

いし、リンクを貼りたいと考えています。

そういった一つのツールができれば非常に役に立つだろうし、中身的に見ても十分にいけるというものであれば、いろいろな医療機関がそこにリンクを貼るという形も考えられます。患者さんがアクセスして医療情報を取りたいと思ったときに、サーチエンジンで探すのではなくて、ある病院の例えば「病気情報の広場」とかといったところから行けるようになると、非常によいのではないかと思います。

恐らく多くの患者さんとそのご家族は大変苦労してたどり着けなくて困っているのではないでしょうか。またこういったページは、実は外来診療でドクターがインフォームドコンセントするときの非常に強力なツールになってくるのではないでしょうか。パンフレットを横に積んでおいてお渡しするよりも、必要に応じて、その方が良いかと思うのです。いかがでしょう。

大矢： この採点ツールをウェブサイトで公開して、それで、医師はもちろんですけれども、患者様もこれにアクセスして、コピーしてそれを見ながら、自分のもらったパンフレットなりウェブサイトを採点して比べていくということができるようにしたいと思います。

会場： 大変重要な結果で、非常に良い研究だと思います。

コメントとして、このDISCERNはどうも相対評価は非常に良いけれども、絶対評価に問題があるということのようですが、ただユーザー側としては恐らく絶対評価として運用するというのが多くのパターンになりそうなので、絶対評価を高めるという意味で、相対評価はちょっと落ちるかもしれませんし、ひょっとしたら信頼性係数がちょっと落ちるかもしれないのですが、3段階評価にするといかがでしょうか。絶対評価としては堅牢になると思います。

信頼性というところで、DISCERNの16項目の項目間のクローンバックα係数は出されましたか。

大矢： それは出しておりません。

会場： そうですか。それが十分に高ければ3段階でも問題ないのではないかと思います。

大矢： 内的整合性は恐らく高いと思います。ただ、κを出すときには一致していないと0になってしまうので、そのへんで低く出たのだと思います。