

法を工夫する必要があるのかと思っております。以上でございます。ありがとうございました。

渡辺：ありがとうございました。以上でございますが、なにか質問はありますでしょうか？

南澤：非常に素晴らしいデータだと思います。ここまで高精度にできるとは思ってもおりませんでした。この虚実の判定はどのような形で？これは慶應のデータですよ？

井元：はい、用いているものは慶應のデータです。慶應で診断された虚実の証を正しい診断であるとして、その証を予測する式を構築しています。

南澤：虚実は何段階の？ややというのがあったと思うのですけれども。

井元：もともと5段階です。予測式は、5段階の中で一番極端な2つ「虚証」と「実証」を比べることで作りました。

南澤：渡辺先生、先ほどの気血水もこれと全く同じですよ。5段階慶應もやっておられる訳ですよ。

渡辺：気血水はうちは5段階やっていないです。

南澤：いや、虚実は5段階やっておられるのですよね。

渡辺：はい。気血水はその中の6つあるのと、最近 ICD に合わせて下焦の虚というのが加わった7つの中から1つないし2つを選ぶというルールにしたのですよね。ICD は木村先生も絡んで一緒にやっているのですけれども、ルールが変わったりしたので、慶應のもともとのデータがどこまで使えるのかというのもあって、新しいプラットフォームでやった方がたぶんデータとしては精度が高いかなと思います。

南澤：ICD と今回の漢方の診断からの支援システムというのは必ずしも全く合致させる必要はないと思うし、ちょっと合致するのは難しいと思うのですよね。だから当然気血水というのは並列のデータですから。

渡辺：そうです。おっしゃるとおりいくつでも選べるのですけれども、最初はいくつでも選べ

るというルールでやっていたのです。まだ論文にしていなくて申し訳ないのですけれども、7、8年前くらいに証のコードのコーディングというのを東洋医学会の理事の先生方、代議員の先生方をお願いしてやったことがあるのです。そうすると気血水に関してある先生は1つだけ選ぶと、ある先生は6つくらい全部選んでしまうという事でコーディングルールがない状態でやると全くデータがとれなかったというのが問題になったので、それで1つまたは2つまでというルールでICDはやろうというふうに東洋医学会の中では話しをしている所です。

南澤：そこで例えば重み付けがしてあれば、最も重いものを2つ選べ、3つ選べと自由に後でできると思うのですけれども、そこで最初からあるなしでやってしまうと結局使えないという事になってしまうのではないかと思うのですよね。

渡辺：それが重み付けという事で4段階のものが入りましたので、それも活用できれば良いなと思っております。

南澤：それは必須に近い形でやっていただければなと思っております。

渡辺：他に何か御意見ありますでしょうか？井元先生、実際に患者さんにフィードバックする時は、虚実のバーの中のどこかで示すのでしょうか？

井元：患者さんにフィードバックすることがよろしいのかどうかということは、また別の問題としてあるかと思いますが、少なくともドクター側にフィードバックする時には、この患者さんは「虚証です」、または「実証です」という離散的なものよりも、スライド2にありますようなレベルゲージのようなものでどの程度虚証なのかを示す方が情報量としても多いですし、我々が使っている手法の特性からもいいのかなと思っています。

南澤：まさしくこれですよね、重み付け。

渡辺：今、患者の症状の変化はグラフ化して患者と医者とが共用できるようになっているのですけれども場合によってはこれも患者さんの画面にフィードバックしちゃってもいいのかなという気はしています。

井元：医師ではない情報学者としての感覚で言いますと、「これは診断に相当する」と言われますちょっと困ってしまうところです。もし可能でしたら、最初にドクター側に表示させてい

ただいて、ドクター側で修正すべき点はドクター側で修正していただくことができると我々情報の人間としては安心感があるかと思いますがいかがでしょう。

南澤：おっしゃるとおりだと思います。

渡辺：では参考にして。

井元：ちょっと臆病者かもしれませんが。

渡辺：とんでもないです。今、とりあえず初診時の診断を予測する、予測するというのをもうちょっと分かり易く言いますと、あくまでも標準が慶應の今の我々の診断を標準としてそれをコンピューターが学んでいるという事ですので、今回お声かけしてもらった施設、という漢方の中で言うのだいたい古方の仲間なのでそう大きく違わないかなと思うのですが、データの量が増える事によってより精度が増してくるというふうに考えております。ではこの先ですね、なぜ経時的な問診をとっているかと言いますと処方支援というのは、例えば高血圧で実証で熱証だと黄連解毒湯とか、ある程度西洋の病名が決まって漢方の証が決まるとその処方提示というか提案みたいのはたぶんシステムとしてはできるんですね。ただそれは理論上のものであって黄連解毒湯を実際にやったら本当に良くなったかどうかという事まで含めた診療支援というものを将来的に構築する為には今の画面でいうと治療効果の予測という事になりますけれども、将来像としてはこのような患者さんに黄連解毒湯をやったら3ヶ月後にこうなりますよというふうな予測ができるというようなものが最終像というふうに考えております。その手始めの一番初めの漢方の診断の所を漢方にあまりなじみがないようなドクターでも分かるというような仕組み作りは今年度中にはできるかなといった所でございます。

南澤：前回お話をさせていただいたと思うのですが、今効果については効果判定は全く入っていないですね。

渡辺：効果判定は入っていないです。

南澤：結局ある患者さんがいてそれに対して我々がどういう処方を選ぶかという所まではできていると思うのですが、この効果はどうであったのかという一項目加わると先生がおっしゃるような事になってくると思うのですが、しかし処方が変わるのは必ずしもうまくいっていない事が多い訳ですよ、良かったらあまりいじらないので。ですから将来

的にシステムを開発するとかで、この前出した処方が例えば効果がどうだったか VAS でどうだったかを入れると簡単な項目 1 つ追加するだけでかなり有用なデータになると思うのですけれどもいかがでしょうか？

渡辺：項目を追加するというのはどういう事でしょうか？

南澤：例えばですね、前回出した処方はどうであったかという主治医の判断項目です。VAS でも何でも良いですし、非常に有効を 100 として全く無効もしくは悪くなったを 0 にしてですとか。

渡辺：なるほど。しかし主治医が当てにならないので患者さんの情報をとるようにしたのですけれども。

南澤：効果判定について言うと、例えばよくある事は黄連解毒湯証については良くなったけれども、先生ついでにこっちも治してほしいなんていう事になると結局不満度としてはあまり変わらなかったりするけれど処方効いている。なぜ転方したのか？という事があると思うのですね。ですから主治医の意見も入れないと。

渡辺：なるほど。ちょっとそれは将来的な課題ですね。どういう形が良いか、先生またちよつとご提案いただければと思います。で、一番最初にですね、まだ数がそれほど揃っていない時点なのですけれども、冷え症、今冷え症の患者さんが一番多いものですから、冷え症の患者さんを集めてざっくりと当帰芍薬散とか、細かい事は云々言わないで漢方治療が有効かどうかという予測をコンピューターがするというやつでやった場合には 90% くらいの確率でコンピューターがこの患者さんの冷えは漢方治療というざっくりとしたもので良くなるか良くなるかというふうな予測がたったというような事なので、それが例えば当帰四逆加呉茱萸生薑湯でどうなるかですとかどんどん細分化されていってしまうと n の数が低くなってしまいますのでまだそこまでは至っていないのですよね。ただ数が集まっていく過程で将来的にはそういった事も可能であろうというふうに考えております。それが個別化医療のエビデンスという事なのでそこまで行けばようやくエビデンスレベルになるのかなというふうに考えております。では、この件はよろしいでしょうか。次は美馬先生からお願いします。

美馬：私の方では植松研究員と共に診断支援のインターフェイス、診断内容の可視化という点を含めて研究を進めさせていただいているという状況で、具体的には患者さんの問診項目を入力としてやってそれが過去に集積されたデータの中でどういう位置づけであろうかというのを

機械学習等で宮野先生、井元先生の方で得られた先ほどのような結果を含めてどのように見せて行くかという方向で研究を進めさせていただいているという状況であります。詳しい内容に関しましては前回の渡辺先生との議論の内容も含めてパワーポイントにさせていただいておりますので植松研究員の方から発表させていただきたいと思います。

植松：(スライド 1) 東京大学の植松です。よろしくお願ひいたします。(スライド 2) ここで我々が担当いたしております分析と可視化についてざっくりと説明させていただきますと、まず問診で得られたデータがデータベースに蓄積させていると。そこで可視化を最終的にするのですけれども、その間に分析というステップがありまして、いろいろ書いてありますが、症状なり問診項目の頻度分析であったり、どのような項目同士が共起、つまり一緒に訴えられる事が多いかといったような分析を行いまして(スライド 3)、次の段階である可視化という所まで辿り着きます(スライド 4)。その場合どのような切り口でデータを見たいのかという事によって、前段階の分析の手法もずいぶん変わってきますので、まず可視化、どのような切り口でどのようにデータを見たいのかという点がこちらとしてはかなり重要になるかと思っています。(スライド 5) こちらとしてはそのように分析をしたデータを使ってこのスライドのように使っただけののかなという想定をしております、検索といっているのはおそらくこの画面では初診というか来院された方のデータとまたその類似、さまざまな角度から見て類似していると思われる方々のデータを見てそこを集計など、類似度、どのような面で近いのかどのような面で異なっているのかというのを分析、そして可視化してそれを先生方が見ていただいて専門的な知識も含めてシステムの方が類似していると出ている所で専門的にどうなのかという所を判断していただくという感じになるかと思っています。(スライド 6) 一番分かり易い例としては、「冷えがあって眠れない」というような方がいらっしゃったとして、そういう方はたくさんいらっしゃる訳なのですけれども、その中で一番理想的には統計処理などをした結果、なんとなくぼんやりと虚証とやや虚証とに分かれてくれると(スライド 7)。しかもこういったグループに分かれた結果をさらに統計処理でグループ分割していくと虚証の中でもさらに類似したいくつかの患者群に分かれるのではないかなというのが一番理想的なシナリオです(スライド 8)。(スライド 9) 今回のお話なのですが、そのような可視化、先生方としてはどのような切り口でどのようなデータを見たのかというような可視化設計ですね、これに関してこの前の渡辺先生とのシステム側での打ち合わせでもかなり話が出たのですが、やはりこちらとしてはかなり必要だなという、どういう切り口が必要なのかというリクエストがこちらとしてはかなり不可欠であるというのと、それなので、こちらの既存の可視化システム、題材は医療関係ではないのですが、可視化システムがありますのでそれを例にいたしまして、このような操作、どのようなデータ整理が可能で、それを基にしてこちらとしては慶應のデータでいうとど

ういう処理が可能かという事をいちお例として見て頂きたいと思っております。それから2番目といたしまして、時経列データの抽象化、つまり患者さんが何度も退院されて、処方されて、改善されたり、ちょっと停滞したりという事を繰り返されている訳ですけれども、そういうものを基にして、連続した時経列のデータをどのように可視化すれば、どのような切り口で先生方にお見せすればその診療なり、診断・処方の助けになるかという所についても、その切り口についてもまたお話いただきたいなとも思っておりますし、またそれによってもデータ処理、スライドでは抽象化と書いていますが、抽象化手法についてもかなり変わってくると思いますが、2項目ともかなりどのような切り口が先生方が必要なのかという所がこちらとしては一番必要なと思っております。(スライド 10)問診情報による、ここで状態と言っているのはその時点での患者さんの状態を可視化する際のこれは案ですが、可視化をするという、これは我々の既存のシステムで、対象となっているのは大学の講義なのでかなり医療データとは違いますが、これを例示とする感じでご想像いただければと思います。この画面の右上の方なのですが、赤い線で囲んである所ですね、それが初診患者、今注目している患者さんの今のデータと過去の患者さんの初診データですね、例えば愁訴などによって類似度を出してやって、画面上の距離で遠い所にある点はあまり似ていない、そして近い所にある点はけっこう類似していると、愁訴などから見てけっこう類似していると思われる患者の状態を表しているという感じですね。で、その下には比較データと書いてありますが、上の方で比較的似ていると思って画面でている、赤い点以外の灰色の点の人達ですね、その人達の状態というか、この人達はどのような状態でどういう診断が下りましたといったような情報が載っているというように考えてください。左側の青い四角の方、その辺によってデータを絞り込む操作をするという事になります。例えば、先ほど井元先生が示していた証予測などや患者と同世代の人だけに絞って類似度を出して欲しいとか、そういう事がしたい、そういう絞り込みが有効なのであれば、やはりそれは実装する事ができますし、逆にどのような属性で絞り込むと右側の画面に有効、効果的な絵と言いますか類似計算が出せるのかというのが我々が先生方の専門知識からどのようなのが見えるのかが知りたい所です。(スライド 11)ひとつ考えられる動作として、特定の症状から類似群を表示して欲しいという事がリクエストされた場合には、属性ごとに、点がいっぱい出ているのではなくて、左下の方に属性ごとの類似群が4つくらい出ておまして、どこかの群を選択いたしますと、ある一部の群だけからまた類似度なり分割して、右上の画面に出す事ができます。(スライド 12)なので、どのような属性によって絞り込んだり、どのような症状によって絞り込みたいのか、どのような絞り込みをすると漢方医学的に有効になるのか、そのようなところを渡辺先生はじめ専門家の先生方にリクエストしていただきたいという所です。(スライド 13)次に診療と時経列、つまり患者さんは来院を重ねている訳ですけれども、そのような何度も来ていらっしゃるというデータをどのように表示するのが良いのか、という事に関してなのです。

れども、まず、これはまだ案の段階なのですけれども、例えば証や症状で似た方の中でさらにある群を選択した時に、その方々の中で結果的に過去の処方の中でどのような薬剤が使われたかという事を右側の括弧の中に出すと、その中で例えばですが、選択した症状の VAS 値が良くなったか、悪くなったかというのは今の効果予測を付けなくてもデータからすぐにとれますので、例えば、選択した症状に注目して次回来た時に VAS が良くなったか悪くなったかというデータはすぐに出す事はできますけれども、実際にそういう事が見れて、意義がある事なのかどうなのかというのがこちらとしては判断がまだできないのですが、そういったようなオペレーションは実装は可能です。例えばですが、右側の選択した中の下の薬剤 A が処方されたという方々で、さらにあまり効かなかった、次回 VAS 値があまり良くならなかったという方々を群として左側の真ん中の類似度を画面に表す事は可能かなとは思いますが。(スライド 14) そのような時系列の可視化を行うのに必要な処理として、分析という事なのですが、時系列データの抽象化という事がありましてそれを今回はマルコフ過程といわれるものとして、状態遷移経列ですね、と思ってモデル化しておりまして、それはよく似た状態を探してどんどん抽象化していくという事になります。患者さんの問診データにいろいろな属性というかがついていますが、そのうちから特徴を抽出しまして、その抽出した特徴同士の間で類似性などを見ていて、それでこの状態と似ているね、という事になりましたら抽象化してしまうという事になって、(スライド 15) 結局グラフ上で(スライド 16) 同じような状態だとして使ってしまうという。乱暴ですがそのようにすることによってある程度類似した状態を扱う事はできるという事ですね。(スライド 17) そのような抽象化を実際のデータ、これはちょっと古いのですが、去年の6月時点でのデータだと思うのですが、その時のデータに対して行った処理例なのですが、左側が処理前で、ちょっと見にくいのですが、一番上から始まってどんどん下にいくに従って時間が経っていくという事になります。初めはバラバラなのですが、先ほどのような処理を繰り返す事によって右側のような状態になるというのを計算する事ができるのですが、この抽象化というのが良い抽象化で、合っているかどうかというのは可視化にして使ってみて皆様に見ただいて、これは納得できるようなデータだなという事にならないと評価もできないという事になりまして、やはりどのような可視化が望ましいのか決めていただくというか、アイデアを頂くのが始めだろうというのが我々としての結論です。(スライド 18) それでですね、今後としては優先的に可視化の設計というのを決めていくとともに実装、これは既に MIMA サーチという可視化エンジンがありますので、それプラスいろいろ可視化ライブラリを使ってやっていくという事になります。あと、そういった可視化のルールを問診システムと接続していくという話ですね。それと可視化設計が決まった後に時系列抽象の話も本格的に動き出すかなというふうに思っております。こちらからはこれで以上です。

渡辺：どうもありがとうございました。美馬先生の方にいろいろお願いしてやっていたのがまだ最近なのでまだできていないのですけれども、今年度の成果としては井元先生から先ほどお話しありましたように、未来予測とか時系列的なものはちょっとまだ先の話だと思うのですよね。ですからそこはアイデアとしては、計画としては報告書にももちろん載っていただければと思うのですけれども、とりあえず見せてほしいのはMIMAサーチを使った診断の分類なのです。例えば虚証、同じ愁訴で来た人がどういう虚実でいうとどういう分布になっているかですとか、前回のミーティングの最後にちょっと話が出たような虚実・寒熱・気血水というようなものが可視化できるという。ですから気血水も全部ひっくるめてというのが複雑であれば、気虚の可視化ですとか。そもそも美馬先生にお願いしているのは、MIMAサーチというのは過去何回かこの会議で見せていただいているので思い出していただければと思うのですけれども、要は患者さんのクラスター、同じような集団を可視化するという。これでやっている事はまさに漢方の証なのですよね。例えばその血圧が高いといっても虚証の人もいれば実証の人もいるのだけれども、そういったものが全部MIMAサーチによってクラスター化してそれが可視化できると。漢方でやっている事は同じ病気であっても証が違えば薬が違ふ、患者さんのグループ化なのですけれども、まさにMIMAサーチの手法というのは2000年の歴史で漢方をやってきたその再現なのです。ですからちょっと時系列のところはまだちょっとハードルが高い所なので、データの数も足りませんので、まずは診断の所をですね、可視化いただければなというように思うのですけれども、何か御意見ありますでしょうか？

南澤：時系列の件ですけれども、これはやはり効果判定を入れないとたぶんデータになりませんよね。

渡辺：今の植松先生のお話だと、効果判定もそうなのですけれども、あるグループの人には過去においてこんなものが出されていますよという事の例示ですよね。それだけでも漢方の専門家の知識を共有する事はできるので、それは診療支援のひとつのツールになると思います。ただそれには数が足りないのです、ちょっと今の時点では難しいかなと思います。

南澤：これはどういうふうにするかだと思うのです。例えば目の前の患者さんにどういうふうにするかというのと、例えば証の解析みたいに使うのか。

渡辺：診療支援と我々が言っているのはですね、なぜこのメンバーでチームを組んでいるかというと、この日本を代表する漢方のプロの集団の知識を一般の開業の先生方と共有すると。いつも学会の中で私が言うのは日本の医師数が今287000くらいで、その9割が漢方を使っている



のですよね。そうすると実に 260000 人くらいは漢方を使っているという事になります。このチームに加わっている皆さんは漢方専門医なのですがすけれども漢方専門医は 2400 名しかいないと。この 260000 対 2400、このギャップをどう埋めるのかという所がこの診療支援なので、決してこれは南澤先生の診療を支援しようと思っている気はさらさらないです。一般の方々、先生方の参考になるような物を作るというのがこのアイディアなのです。

南澤：この可視化の目的を実際の目の前の患者さんに対して役に立つように見せて頂くのか、それとも視聴というのを全体を、母数を、何千人という患者さん達を検討に使うのか。

渡辺：目の前にいる患者さんの入力を見て、それをコンピューターが解析して、例えば頭痛で来た人と、頭痛で来た患者さんはいくつかのグループに分けられるのだけれども、目の前の患者さんは頭痛の中でも虚証で瘀血のあるグループに入りますというような事が可視化できるというアイディアです。

南澤：なるほど。では時系列というのは先ほどの効果をまた加味したデータとしてやっていくという事ですよね。

渡辺：時系列は、すみません、最初に私が美馬先生に時系列を加味した可視化と言うことでお願いしたので、たぶんそれにお答えしようとしてくださっているのだと思いますけれども、ちょっとまだこの時点ではデータの数も足りませんし、ハードルが高いかと思うので。

南澤：たぶん効果判定をいれないと時系列はあまり有効なデータにならないという事ですよね。

美馬：おっしゃるとおりで、やはり我々も先ほどお示しさせていただいた様に、患者の途中の経路が似ている所をどんどん融合してやって抽象化と、要するに同じような状態として認識できるようにしていくのですけれども、その過程で同じ状態という事だけではなくて、例えばある葛根湯なりを投薬された時にどうなったのかというのは、先生がおっしゃたように効果ですね、それがどういように判断すればよいのかというのは一つの我々としての問題にもなっております。要するにある投薬をして、当然効く、定量的にいうと VAS 値があがる、あがるかどうかは微妙ですが、変化があると、要するに良くなるグループと悪くなるグループ、悪くなるかよく分からないですけれども、要するに変わらないグループ等に分かれる訳ですけれども。そうすると分かれた先を同じように併合していくという事はできない訳ですよね。そういう意味で、通常マルコニフデシジョンモデルというのに我々は落としていこうとしているので

すけれども、考えた場合にはある状態から、ある投薬をして、次にどうなったかという、その3点セットで入力をしてやって、最終的には例えば今の初診の状態からある投薬をしたらどれくらいの確率で良くなるかというその確率値がでるわけですね。まずは何か良い状態になったという事を判断できるのかという話とそこからさらに我々が今やろうとしているのは確率なりがでた時にそれをどう示すか、どう可視化するかという話ですね。そこがやはりこういう議論の中でトライアルをしながら良くしていくというステップになるかなと思っている所です。

渡辺：他にいかがでしょうか？よろしいでしょうか？美馬先生、ちょっと時系列は次年度以降という事でまずは診断の所の可視化を是非ともお願いいたします。それでは、議案の4番、平成23年度の報告書という事で、去年は問診の各施設の比較のようなもので、書いていただいたのですけれども、今日は御参加の中では山本先生と南澤先生は研究協力員なので該当しないのですけれども、後の施設の先生方には頭の痛い報告書というものがございます。期日は5月の連休明けなので連休明けまでになのですけれども。去年、今年の研究がシステムの導入が少し遅れたこともあって少しご迷惑をかけているかなとも思うのですけれども、まず富山から、だいたいこんな構想とかいうのはございますでしょうか？

引網：初年度にどういうふうにするかというのは2つありまして、1つは前年度の内容なのですが、もうひとつは従来問診票、我々の施設でやっていた問診票の過去数年分の解析をして方剤との適応証というのをやろうという事で今やっているものを報告したと思っております。

渡辺：それでよろしいかと思えます。そういう事でお願いいたします。あと経過としては倫理委員会を通してこれを導入する準備ができたというような所までかと思えますけれども、ではそれをお願いいたします。それでは、女子医大お願いいたします。

木村：うちの方では元々うちが持っているシステムを倫理委員会に出すのが夏くらいに通ったので、それを受けての今度慶應の事という事で今やっている所です。それで、ちょっと報告書についてはこの前機械をいただいて、その前にいちお問診票の初診の方の入力などもしているのでそれから何か、去年は頭痛だったのですが、何か主訴が多いので報告書にしたいと思っております。

渡辺：分かりました。ではそれをお願いいたします。トムラスとの比較でも良いですね。

木村：そうですね。ただ入力システムが今週月曜日に頂いて、それから5月までにデータが

集まればですけども。今トムラスが去年のシステムでいただいたので合わせて入力しているので、間に合えばそういう事もしてみたいなと思っているのですが、そこらへんは件数を見ながらさせていただきたいと思っています。

渡辺：分かりました。予算の執行の方は大丈夫でしょうか？

木村：予算の方は、入力で半分くらい使わせていただいております、あとは備品なので、という感じです。

渡辺：大学の規定、たぶん施設によって違いますが、2月の末くらいまでです。

木村：入力の人件費でかなり使わせていただいておりますので大丈夫です。

渡辺：はい分かりました。並木先生聞こえますか？

宗形：音声は聞こえていらっしゃいますが、マイクが使えないと連絡いただいております。

渡辺：では後でメールでやりとりという形になりますでしょうか。村松先生は？

宗形：抜けられています。

渡辺：では、井元先生の医科研と東大の方は今のような事をまとめていただくという事で。

宗形：並木先生よりコメントが入りました。

渡辺：「慶應の問診と気血水スコアの比較、慶應の新しい問診項目への追加の経緯を書きます」と。はい、ありがとうございます。ではそれをお願いいたします。報告書に関してはまたメールでやりとりをさせていただきたいと存じます。その他としましては平成24年度の申請書は先生方に送らせていただいたと思います。中間審査というのが、今年の3月1日にございますので、そこでiPadに載ったというようなwebのやつと、井元先生、美馬先生の可視化の方も含めて構築をして発表したいと思っています。それからもうひとつのアナウンスとしてはこの問診のシステムについてNHKの「名医にQ」という番組の福澤先生が出られる番組の中で紹介というので、バタバタと先週話があつて収録に来て、なんと私の診断と予測がちゃんと合ってい

るかというような事をブラインドでやれという事で、先週の土曜日に収録をして、解析を井元先生にお願いしたのですが、虚実の中間証で、寒熱は錯雑という方だったので、見事に一致しましたので。

井元：少しコメントさせていただきますと、その患者さんは、先ほど渡辺先生からお聞きした所によると本当は患者さんではなく、元気な方だったという事でした。問診データについても我々は、20項目以上入力がある患者さんを対象にしていますが、実はこの方は入力されている問診項目が22～23項目でぎりぎり解析できるレベルでした。結局、虚実は中間証と予測しました。先ほど説明しましたように、一番難しい予測です。また、寒熱は錯雑証。これも一番難しいところですよ。一番来て欲しくなかったテストデータでした。それでも幸いな事に予測が当たっていたという事で、我々は本当にほっとしています。

渡辺：いちお放映日は2月18日ですが、地震の特集が何かで延びるかもということなので、まだ未定のようです。

南澤：決定したらぜひ教えてください。

木村：なんていう番組名ですか？

渡辺：「きょうの健康」の「名医にQ」という、「きょうの健康」の長い版だと思えるのですが、だいたい以上で用意した議案は終わりなので、それで、来年もう1年度続くという事で各施設で集めたデータで各施設間のデータの比較なんかできますし。それから今日はそこまで踏み込んで話しができなかったのですが、問診の項目も引網先生からご指摘のように、ちょっと高齢の方に優しくないとか、項目が多いと。要は虚実、寒熱、気血水、六病位の診断ができればだいたい目的が果たせるので、その予測に使う項目、使わない項目というのが、今井元先生の解析の方で明らかになりつつあります。例えば、慶應の項目で言いますと、去年の和漢医薬学会の発表でもあったのですが、右足の冷えと左足の冷えを区別する必要があるのかどうかという事で、要するにリンクして動くようなもの、リンクして〇がつくようなものは省けるのではないかと。もうひとつは虚実・寒熱とかの予測の使うのが非常に貢献度というのですかね、ちょっと井元先生の方から補足していただきたいと思うのですが、予測に使う寄与率が極めて低いものがあると。虚実の診断に使う寄与率が低い、寒熱も低い、気血水も低いといった場合には、全てのものにおいて寄与率が低いというものには、そういったものを切り捨てていくという事も可能ではないかという議論が出てお

ります。ちょっと井元先生補足をお願いいたします。

井元：あまり補足させて頂く事はないのですが、寄与率というのは先ほど発表しました内容では重要度と申し上げたものに相当します。虚実の予測では、BMI が1番重要度が高いことが分かりましたとお話した事に相当します。今の所、虚実と寒熱の予測に関してこの重要度が各問診項目について得られています。これを気血水に関しても計算いたしまして、この重要度が低いものに関してはリストラではないですが、削除の対象となるであろうと考えています。ただし、証の予測における重要度は低いけれども、何かしら診察・治療の助けになるものはあるでしょうから、その点は皆さんの議論を待ってというような事になるかと思えます。

渡辺：要するに、この予測に使っているのは 200 項目でしたっけ？すべて使っているのですか？

井元：すべて使っています。すべてに関して重要度が計算されます。

渡辺：それで BMI に関して寄与度が 3 くらいでしたか、一番低いやつが 0.00 何とかという感じなので、ほとんど寄与がないものもあるという事で、少なくとも問診項目 200 の中の BEST200 全部が出るという中で、あまりにも寄与率が低いものはカットできるかという議論というものは当然可能かと思えます。そういう訳で患者さんにもっと使いやすいものにどんどん改良は続けていくという形を取りたいと思います。他に何かございますでしょうか？それでは引き続き研究の継続をお願いいたします。それから協力研究員の南澤先生、山本先生、今日は参加されていないけれども田原先生も含めてですね、データの補助が必要だとか予算の措置というものはご相談いただければ個別に対応したいと思います。では、他になれば事務局から何かアナウンスありますか？

宗形：ありません。

渡辺：ではこれで終わりたいと思います。どうも長時間ありがとうございました。

南澤：渡辺先生、具体的にはメールか何かさせていただけばよろしいのですか？

渡辺：そうですね。もし今年度から始めたいのであれば何かのお金は融通できますけれども。

南澤：どういうふうに掲載するかとかその辺もあるかとは思いますが、滑りだしはちょっとなんとかやり繰りしてみようかとは思っていますので。ただうちは科としてのお金は全然ないので。

渡辺：それはたぶんできると思うので、慶應ルールで例えば慶應がお金を払った人が外で働くのがよいのかどうかとか、そういう抜け道をちょっと模索しないとですね。相談していただいでうちの中の研究支援センターというのがあるので、たぶん亀田もありますよね。

南澤：そうですね、たぶんあると思います。

渡辺：少なくともお金を管理する所はあるはずなので、そこでの話し合いの中で模索というような事になると思います。始める時期が、3月、2月から始めたいという事であれば十分相談に乗れると思います。

南澤：分かりました。ありがとうございます。

#### IV. 研究成果の刊行に関する一覧表

## 研究成果の刊行に関する一覧表

雑誌

発表者氏名	論文タイトル名	発表誌名	巻号	ページ	出版年
Plotnikoff GA, Watanabe K, Torkelson C, La Valleur J, Radosevich DM	The TU-025 keishibukuryogan clinical trial for hot flash management in postmenopausal women: results and lessons for future research.	Menopause	18	886-92.	2011
Watanabe K, Zhang X, Choi SH	Asian medicine: a way to compare data.	Nature	482	162	2012
Cameron S, Reissenweber H, Watanabe K	Asian medicine: Japan's paradigm.	Nature	482	35	2012
Gao PF, Watanabe K	Introduction of the World Health Organization project of the International Classification of Traditional Medicine.	中 西 医 結 合 学 報	9	1161-1164	2011
Katayama K, Yamaguchi R, Imoto S, Matsuura K, WatanabeK, Miyano S	Clustering for Visual Analogue Scale Data in Symbolic Data An alysis.	Procedia Co mputer Scie nce	6	370-374	2011
Katayama K, Yamaguchi R, Imoto S, Matsuura K, WatanabeK, Miyano S	Transform of visual analogue scale data and their clustering.	International Journal of Knowledge Engineering and Soft Da ta Paradigm s	3(2)	143-151	2011
Mima H, Masuda K, Ota S, Yoshimi S	Revealing Modern H istory of Japanese P hilosophy Using Nat ural Language Proce ssing and Visualizati on.	Proceedings of Osaka Sy mposium on Digital Hum anities(OSD H)		47-50	2011



発表者氏名	論文タイトル名	発表誌名	巻号	ページ	出版年
美馬秀樹	生命科学における知の 構造化	Biophilia	7(3)	47-53	2011
渡辺賢治	漢方医学をめぐる最近 の動向	医学のあゆみ	240	988-990	2012
上野眞二, 村松慎一	Alzheimer病と漢方薬	神経内科	76(2)	147-154	2012

く V. 研究成果の刊行物・別冊

# The TU-025 keishibukuryogan clinical trial for hot flash management in postmenopausal women: results and lessons for future research

Gregory A. Plotnikoff, MD, MTS, FACP,<sup>1,2,3</sup> Kenji Watanabe, MD, PhD,<sup>2</sup> Carolyn Torkelson, MD, MS,<sup>3</sup> June La Valleur, MD,<sup>3</sup> and David M. Radosovich, PhD<sup>3</sup>

## Abstract

**Objective:** The aim of this study was to assess the efficacy of TU-025, keishibukuryogan, a Japanese prescription herbal medicine used for hot flash management, in American women.

**Methods:** This randomized, double-blind, placebo-controlled, phase II trial enrolled 178 postmenopausal women aged 45 to 58 years with a Mayo hot flash score greater than 28 per week who met other inclusion criteria. After a 1-week placebo run-in period, participants were randomly assigned placebo, or 7.5 g/day, or 12.5 g/day groups, for 12 weeks. Primary and secondary outcomes were measured using the Mayo Clinic Hot Flash Diary, the Greene Climacteric Index, and the Pittsburgh Sleep Quality Index.

**Results:** At 3 months, hot flash scores, climacteric symptoms, and sleep quality improved by 34% in the placebo group, 40% in the 7.5 g/day group, and 38% in the 12.5 g/day group. ( $P < 0.001$ ). However, the differences in changes between groups were not statistically significant ( $P = 0.990$ ). Diarrhea unexpectedly developed in 20% of participants receiving active medication.

**Conclusions:** For American women, unlike the clinical experience for Japanese women, TU-025 did not significantly reduce the frequency and severity of hot flash symptoms, improve climacteric symptoms, or benefit sleep quality. This study identified several potentially significant methodological factors to be considered in future scientific assessments of traditional Asian medicines.

**Key Words:** Herbal medicine – Hot flash – Menopause – Alternative medicine.

Postmenopausal hormone therapy has well-documented positive effects on vasomotor symptoms, sleep quality, and even sexual functioning. However, many postmenopausal women are concerned about the US Food and Drug Administration (FDA) boxed warning regarding a slight but statistically significant increased risk of substantial harms including dementia,<sup>1,2</sup> breast cancer,<sup>3-5</sup> endometrial cancer,<sup>6</sup> venous thromboembolism,<sup>7</sup> and gallbladder disease.<sup>8</sup>

The 2010 North American Menopause Society position statement on postmenopausal hormone therapy addresses these concerns and, based on clinical data, supports the use of hormone therapy during menopause to treat significant symptoms

or reduce significant risk of fractures.<sup>9</sup> The Society notes that the potential therapeutic benefits and risks needed to be considered. For symptomatic women concerned about a potentially increased risk of atypical ductal hyperplasia,<sup>10</sup> estrogen receptor-positive breast cancer,<sup>11</sup> and node-positive breast cancer,<sup>12</sup> there is great interest in nonhormonal options.

In Japan, more than 90% of Japanese gynecologists prescribe traditional Japanese multiherb medicinal formulas for women's health concerns including menopause.<sup>13,14</sup> Termed Kampo, these formulas are regulated as prescription pharmaceutical drugs by the Japanese Ministry of Health and have been covered by the national healthcare plan for more than 25 years.

The 1,800-year-old formula termed keishibukuryogan, whether known as TJ-25 (the Japanese identifier) or TU-025 (the US identifier), is one of the foremost Kampo agents used for perimenopausal hot flash management in Japan. Keishibukuryogan consists of cinnamon bark (*Cinnamomum cassia* Blume), peony root (*Peonia lactiflora* Palls), peach kernel (*Prunus persica* Batsch), poria sclerotium (*Poria cocos* Wolf), and moutan bark (*Peonia suffuticosa* Andrews). The active ingredient(s) is not known. Synergistic interactions are presumed based on historical record and experience.

Kampo agents meet the stringent Japanese standards for product manufacturing, consistency, and stability. Postmarketing surveillance data obtained from the Japanese Ministry of Health and Welfare and Tsumura & Company demonstrated

Received October 27, 2010; revised and accepted December 22, 2010.

From the <sup>1</sup>Allina Center for Health Care Innovation, Minneapolis, MN; <sup>2</sup>Keio University School of Medicine, Center for Kampo Medicine, Tokyo, Japan; and <sup>3</sup>University of Minnesota Medical School, Minneapolis, MN.

Funding/support: This work was supported by Tsumura & Company and National Institutes of Health Grant M01-RR00400.

Financial disclosure/conflicts of interest: Drs. Plotnikoff and Watanabe have received funding from Tsumura & Company for research, presentations, and consultations. Drs. Plotnikoff and Watanabe have received honorariums from Tsumura & Company for speaking engagements. The other authors report no conflicts of interest or need for any disclosures.

ClinicalTrials.gov identifier NCT00119418.

Address correspondence to: Gregory A. Plotnikoff, MD, MTS, FACP, Allina Center for Health Care Innovation, Abbott Northwestern Hospital, Mail Route 33540, 800 E. 28th Street, Minneapolis, MN 55407-3799. E-mail: Gregory.plotnikoff@allina.com

that TJ-25 used in Japan has an estimated adverse event rate of 0.00356% for the 128,435,263 doses provided to 701,832 unique women treated for an average of 6 months during a period of 9 years.

Studies conducted using state-of-the-art approaches have demonstrated that TJ-25 has no estrogenic activity. Noguchi et al<sup>15,16</sup> in Japan have developed an animal model of menopausal hot flashes focusing on calcitonin gene-related peptide (CGRP). Animal model in vivo interventions of TJ-25 result in an equivalence of TJ-25 compared with estrogen in suppressing CGRP-induced hot flashes with no evidence of estrogenic activity.<sup>17</sup> In addition, in the same animal model, TJ-25 inhibits the synthesis and release of CGRP.<sup>18</sup> Given these mechanistic considerations, we conducted a randomized, double-blind, placebo-controlled, modified phase II trial that enrolled healthy postmenopausal American women to determine the impact of TU-25 on postmenopausal symptoms.

## METHODS

### Design

The study was of 12-weeks duration plus a 1-week placebo lead-in period. Postmenopausal was defined as (1) no menses for consecutive 12 months, or (2) previous hysterectomy without oophorectomy and follicle-stimulating hormone (FSH) levels of 40 mIU/mL or higher and estradiol levels of 20 pg/mL or lower, or (3) 2 months past surgery for hysterectomy with oophorectomy. Inclusion criteria included vasomotor symptoms documented by the Mayo Hot Flash Symptom Diary scoring system: mild (1), moderate (2), severe (3), and very severe (4), for a total score of 28 or greater in a week (ie, two moderately severe hot flashes/d =  $14/\text{wk} \times 2 = 28$  score or 4 mild/d =  $4 \times 7 = 28$ ). Other criteria included willingness to take an herbal formula, no hormone therapy for 8 weeks, smoking of less than 10 cigarettes/day, body mass index (BMI) less than 36 kg/m<sup>2</sup>, Beck Depression Inventory score of less than 11, and not currently taking antidepressant medication. Women with a history of breast or uterine cancer were excluded; however, for other cancers, individuals were eligible if they were cancer-free for 5 years.

No pharmacokinetic data on the multiple potentially active ingredients in keishibukuryogan exist. For this reason, regular dosing was defined as the dose amount provided on a prescription basis to women in Japan (7.5 g/d). To assess for a possible dose-response in larger-sized American women, we included a higher-dose arm of 12.5 g/day. Participants were block randomized to receive one of three treatment protocols: (1) TU-025 at 7.5 g/day, (2) TU-025 at 12.5 g/day, or (3) placebo. Five tablets of the product were taken twice a day for 12 weeks. Participants were seen monthly for compliance, adverse effects, laboratory monitoring, and outcome measures. Tsumura & Company (Tokyo, Japan) provided the product and placebo in a capsule form. A separate research pharmacy service distributed placebo and active medication that were identical in appearance and odor. All participants took placebo for 1 week after enrollment before active medication was provided.

The study was approved by the University of Minnesota Research Subjects' Protection Program. All methodologies were reviewed and approved by the FDA under their Investigational New Drug policies. The study was registered at Clinical-Trials.gov (NCT00119418).

### Sample

From more than 2,200 inquiries, 178 healthy postmenopausal women were recruited and enrolled from the Minneapolis, St. Paul metropolitan area, through newspaper, radio, and television coverage.

### Procedure

The women were randomized in blocks using a pseudo-random number that was used to assign the participants, as they were enrolled in the study, to the placebo, low-dose, or high-dose groups. The randomization schedule was sent to the research pharmacy to prepare containers for administration to the study participants. All investigators and research staff were blinded to the group assignment. The randomization and all analyses were conducted using SAS System software.

Analysis followed the intention-to-treat guidelines. Data were analyzed using general linear mixed model methods and the generalized linear mixed model methods approach.<sup>19-21</sup> The outcomes were analyzed as either a continuous variable or a categorical variable when there were clinically accepted cut-points for the measurement. For the continuous variable outcomes (the hot flash scale score, the Greene Climacteric Index [GCI], and the Pittsburgh Sleep Quality Scale [PSQI]), we used the PROC MIXED procedure in the SAS System software. For categorical outcomes (hot flash scale score  $\geq 28$  points), the PROC GENMOD and GLIMMIX procedures were used.

For efficacy analysis, the average daily hot flash score was weighted based on the number of mild symptoms plus twice the number of moderate hot flashes plus three times the number of severe hot flashes plus four times the number of very severe hot flashes recorded in a given week divided by the number of days. For the primary efficacy analysis, changes in the hot flash score and the number of hot flashes were summarized as the change in percentage from the baseline.

The type 3 test of fixed effect provided the F statistic, degrees of freedom, and *P* values for the effects of the treatment group, month, and treatment group by month. The difference of differences in the means was tested using the latter interaction term.

Statistical power analysis was based on the published data of Loprinzi et al<sup>22</sup> on self-reported frequency and severity of hot flash symptoms with interventions. For the within-group SD for Composite Hot-flash Score, we used the SEM and sample size for the placebo group from the paroxetine randomized controlled trial by Stearns et al.<sup>23</sup> Based on their data, we estimated the SD for the Composite Hot-flash Score as 8.76. For statistical power, we assumed a 5% level of significance and a 90% power for detecting a true group by level difference. We assumed 10% attrition in the study. Based on these assumptions, the study was adequately powered with 60 enrolled participants for each of the three study arms.