

A. 研究目的

Evidence Based Medicine (EBM)が提唱されてから約 20 年が経過し、エビデンスに基づいた医療行為の実践は多くの医療者に浸透しつつある。しかしながら、これまでの我が国における医学分野における研究は基礎実験研究が主流で、臨床研究に関する認識や理解度が十分でなかった^{1,2)}。ここ十数年で徐々に臨床研究や疫学の知識や技術を備えた専門家がいる大学病院や基幹病院に所属する医療者が自ら企画する臨床研究を実施しはじめたものの、プライマリ・ケアに従事する医療者が自ら臨床研究を計画、実践し、エビデンスを世界に発信するまでに至る例は未だ多くはない。プライマリ・ケアというある種各専門領域の診療の疑問と比べ、複雑な診療に直結した疑問から臨床研究を計画し実践するには、日常プライマリ・ケア診療を行っている医療者こそが臨床研究を実践し、その成果をエビデンスとして自らの患者に還元することすることが重要である。

筆者らは、2009 年より厚生労働科学研究臨床研究基盤整備推進研究事業「実現・持続可能性ある臨床研究フェローシップ構築に関する研究」班で医療者を対象にした臨床研究のリテラシー教育および自らのリサーチ・クエスチョンに基づいた研究プロジェクトの立案・実施・論文化のプロセスを通じて学習させるオンザジョブ・トレーニング等を目的として立ち上げた。今回その一環として、プライマリ・ケア診療を行っている医療者の臨床研究者の育成を目的とした臨床研究教育ワークショップ（以下、本ワークショップ）を実施し、教育的効果の評価を行ったのでその結果を報告する。

B. 研究方法

1. 本ワークショップの概略

今回我々が企画した本ワークショップは、参加者が臨床研究の知識やスキルをより効果的に習得できるようレクチャー、小グループ実習およびグループ発表会の構成でプログラムの編成を行った(図 1)。

レクチャーは、本ワークショップで取り上げる臨床研究の基礎知識のレビューに加え、参加者との対話形式の講義を実施した。

小グループ実習は、3~4 人の参加者に対し 1 名のチューターを配置するグループで実施した。小グループ実習の課題は、シナリオから考えられる臨床の疑問や問題点を洗い出し、その疑問や問題点に対する改善をグループで討論させた。具体的には、事前に参加者に提示し検討させた提示した“外来診療待ち時間”をテーマとしたシナリオを、リサーチ・クエスチョン(P:患者、E(I):要因(介入)、C:対象、O:アウトカム、略して PE(I)CO と呼ばれる枠組み)に構造化し、構造化したリサーチ・クエスチョンに対する吟味を行う実習を実施した。実習の各作業には専用のワークシートを用いて実施した。

グループ発表会は、全グループの作業結果を 1 グループ 5 分でレビューし、これに対して他のグループ、講師及びファシリテーターからの質疑応答や議論フィードバックを行った。また、参加者およびスタッフ全員が投票によって優秀賞の選考を実施した。

2. 本ワークショップの評価

1) 本ワークショップ参加者の学習効果の測定

参加者の学習効果を測定する目的で、本ワークショップの学習テーマである「現場の疑問を、臨床研究で回答可能なリサーチ・クエスチ

「オンに構造化する」に関する理解度テスト(以下、理解度テスト)を行った。理解度テストは、Web 形式による方法で行い、本ワークショップの開講 2 週間前(プレテスト)と本ワークショップ終了 1 週間後(ポストテスト)の 2 回実施した。この理解度テストは、臨床研究における研究課題(リサーチ・クエスト)を主要な要素である PE(I)CO に構造化することに関する知識やスキルを問う問題を実施した。ポストの理解度テスト問題内容は、プレの設問を並べ替えたものとし同一問題とした(表 1)。

2) 本ワークショップ参加者によるプログラム評価

参加者によるプログラムを評価する目的で質問票調査を行った。調査は、Web 形式による方法で行いポストテスト時に実施した。調査項目は、本ワークショップの全体評価及び各項目(ワークショップのレクチャー、小グループ実習およびグループ発表会)に対する評価とし、「とても満足している」から「全く満足していない」の 5 段階のリッカートスケールで回答する方式とした。さらに参加者の自由意見の記載を依頼した。

また、参加者の属性として、性別、職種、実務経験年数、臨床研究の経験の有無についても尋ねた。

3. 統計解析

統計解析は、IBM SPSS Statistics 20.0 を用いた。理解度テストの正答割合の前後比較は Fisher's Exact 検定を用い、正答数の前後比較については、対応のある t 検定を用いた。有意水準は 5%未満を有意差ありとした。

4. 倫理的配慮

質問票調査や理解度テストへの回答を参加者に依頼する際、調査結果や理解度テストの成績等の個人情報、個人が特定される

形では公表されないこと、本研究以外の目的に使用しないことを説明し、調査協力への同意を得た。

C. 研究結果

本ワークショップは、2011 年 7 月 2 日及び 3 日に第 2 回日本プライマリ・ケア連合学会学術大会(札幌)にて「臨床研究デザイン道場」として行われた。参加者定員 45 名は事前申込みにて定員に達した。しかし、本ワークショップと学会の一般演題発表時間等が重複するなどの理由で、申込者のうち 8 人が参加困難となり、最終的な当日参加者は 37 人となった。

参加者 37 人の内訳は、男性が 28 人(75.6%)、女性が 9 人(24.4%)であった。職種の内訳は、医師が 26 名(70.3%)、薬剤師が 7 名(18.9%)、看護師が 1 名(2.7%)、鍼灸師が 3 名(8.1%)であった。臨床経験年数は、2 年未満が 2 名(5.4%)、2 年以上 5 年未満が 4 名(10.8%)、5 年以上 10 年未満が 16 名(43.2%)、10 年以上 20 年未満が 6 名(16.2%)、20 年以上が 4 名(10.8%)、回答なし 5 名(13.5%)であった。臨床研究を自ら行ったことがある参加者は 15 名(40.5%)を占めた。

理解度テスト及びアンケートの質問票調査の有効回答数(有効回答割合)は、プレ 32 人(86.5%)、ポスト 20 人(54.1%)であった。

1. 本ワークショップの評価

1) 本ワークショップ参加者の学習効果の測定

理解度テストの正答割合(%)および全設問の正答数(平均正答数±標準偏差)の集計結果を表 2 に示す。表 2 にみられるように、プレの各設問の正答率(%)は「設問 4」の 37.5%から「設問 2」及び「設問 5」の 84.4%、ポストでは「設問 4」の 40.0%から「設問 2」の 100.0%と、大きなばらつきがあった。本ワークショップ前後

の比較では、いずれの設問においても正答割合は増加し、最も正答割合が増加した設問は、「設問 1」で 46.4%増加した。また、全設問の正答数(平均正答数±標準偏差)は、プレでは 3.0±1.0 問、ポストでは 3.9±1.1 問と 0.9 問上昇した($p<0.01$)。

2) 参加者によるプログラム評価

本ワークショップの全体評価及び各プログラム(ワークショップのレクチャー、小グループ実習およびグループ発表会)に対する評価に対する評価結果を表 3 に示す。表 3 にみられるように全体評価及び各プログラムともに「とても満足している」、「満足している」回答した参加者が大部分を占めた。また、参加者の自由意見を表 4 に示す。

D. 考察

本ワークショップは、筆者らによるこれまでに開催した様々な専門医や職種に対する臨床研究教育ワークショップへの反省に基づいて、プログラム内容や方略に改善や工夫を加え、新たに立案したものである。公募後、参加申し込みが盛況で、早々に定員を満了したことから、プライマリ・ケアに従事する医療者の臨床研究の学習に対する高いニーズが確認された。

参加者の学習効果の分析結果から、理解度テストの全設問の正答数が本ワークショップ前後の比較において統計学的に有意に上昇したこと(表 2)から本ワークショップが一定の学習効果をもたらしたことを示唆している。なお、理解度テストの「設問 4」については、プレ、ポストいずれにおいても正答割合が他の設問と比べ低かった。これらの項目は参加者にとって理解が困難な部分である可能性が高く、レクチャーでの詳細な解説や小グループ実

習で議論を行う課題のひとつに取り入れるなど、今後の方略の改善のための参考となった。

次に、参加者による全体に対する評価及び各プログラム評価について考察する。両評価ともに、非常に高い評価を得た理由として、本ワークショップではレクチャーに加え、小グループ実習および発表会を取り入れたことが挙げられる。一つのテーマについてグループで議論することにより、思いもかけない他者の意見が聞けたり、議論する中で思考や方向性が集束していく経験をすることの重要性の認知、そしてこれらの作業を通じて理解が深まったと考えられる。同様の意見が自由回答にみられるように、参加者のニーズに合ったワークショップであったと考える。

一方で、小グループ実習や発表会に時間をかけてほしかったという意見もみられ、今後の改善点としたい。

本調査は、参加者の属性が、臨床経験が 5～10 年の医師が多いことから、臨床現場で勤務する全てのプライマリ・ケアの医療者の意見を反映するものではなく、結論の一般化には慎重を要する。また、今回のワークショップの学習内容は、広範な臨床研究に関する知識やスキルの一部のみを扱ったものであり、理解度テストもその学習領域のみの理解度を測定したものである。これに加え、本ワークショップ参加者が今回得た知識やスキルを基にさらに学習を進め、深めたかどうか、実際に臨床研究を計画・実施し、学会や学術論文で成果報告しているかなどについても調査する必要がある。すなわち、ワークショップのみの学習効果のみならず、臨床研究の継続的学習や実施にいかなる影響を与えたかも今後評価が必要であろう。

以上のような検討課題が残るものの、これまでプライマリ・ケアの医療者のために考案された臨床研究教育ワークショップを実施し、その教育的評価を行った報告は本邦にはなく、我々の立案したワークショップは短時間でもかなりの学習効果や満足度を達成することが可能であることを示した。今後ますます重要視されるプライマリ・ケアにおいて現場で診療業務に携わるプライマリ・ケアの医療者に対して臨床研究の知識やスキル提供していくことこそが、地域医療の充実に繋がる可能性がある。よって、このような単発的な臨床研究教育ワークショップにとどまらず、臨床研究のより系統的・継続的な教育、さらには臨床研究を実施の実現が期待される。

なし

2. 学会発表

なし

G. 知的財産権の出願・登録状況(予定を含む)

1. 特許取得 なし

2. 実用新案登録 なし

3. その他 なし

E. 結論

プライマリ・ケアに従事する医療者を対象とした臨床研究教育に関するワークショップの実施とその評価を行った結果、本ワークショップは一定の教育効果や評価が得られ、プライマリ・ケアに従事する医療者のニーズに呼応した教育プログラムである可能性が示唆された。

(参考文献)

1)福原俊一:臨床研究者育成のための戦略とロード・マップ、学術の動向、8、18-22、(2006)

2)福原俊一:Population-based Medicine の教育 個人から集団へエビデンスをつくる臨床研究者育成—新しいリサーチコミュニティの創生、医学教育、38、83-88、(2007)

F. 研究発表

1. 論文発表

表 1. 臨床研究教育ワークショップの評価結果

	『リサーチ・クエスチョンの作り方』 の講義を理解した		『学会誌編集委員の立場から』 の講義を理解した		グループワークは学習効果を 高めた		今回のワークショップは満足し た	
	人数	%	人数	%	人数	%	人数	%
とてもそう思う	2	10.0	3	15.0	13	65.0	8	40.0
ある程度そう思う	16	80.0	14	70.0	7	35.0	11	55.0
あまりそう思わない	0	0.0	0	0.0	0	0.0	0	0.0
全くそう思わない	0	0.0	0	0.0	0	0.0	0	0.0
どちらとも言えない	2	10.0	3	15.0	0	0.0	1	5.0

表 2. 「疑問をリサーチ・クエスチョンに構造化する」に関する理解度テストの正答率と全設問の正答数

	プレ N=32	ポスト N=20	<i>p</i> -Value
設問1の正答率 (%)	50.0	75	0.07
設問2の正答率 (%)	84.4	100	0.08
設問3の正答率 (%)	40.6	85	< 0.01 **
設問4の正答率 (%)	37.5	40	0.55
設問5の正答率 (%)	84.4	90	0.45
全設問の正答数(問) (平均±S.D)	3.0±1.1	3.9±1.1	< 0.01 *

Fisher's exact test **: $p < 0.05$, Paired student's *t*-test *: $p < 0.05$

表 3. 臨床研究教育ワークショップに対する自由意見

グループワークに対する意見	
とてもよかったです。 研究は一人ではできないな、と感じましたが。	
人数が少なかったが、ファシリテーターの先生のお陰で理解が深まった。	
医師、薬剤師、立場の異なる医療職がひとつの医療問題に取り組むことはとても良かったが、時間が短く、形式的な話し合いとなった部分もあったと思う。	
グループでPECOを考えることで、自分だけでは考えられないようなことまで知ることができた。 また、思ったことや考えたことを議論すると、高め合うことができると思う。	
時間は若干少なめだったが、ファシリテーターの竹上先生にもアドバイスいただきながら現実的に必要性の高い結果についてRQをたてることができ、非常に勉強になった。	
もう少し議論時間が欲しかった	
ワークショップの際は御指導ありがとうございました。 内容的には、非常に実践的で、かつ楽しめたのですが、時間が少し足りなかった点が残念でした。 また様々な背景の参加者とコミュニケーションをとることで、気付かなかった点をお互いに指摘することができ、新鮮でしたが、なかなか意見がまとまらず、少し大変でした。	
短時間で曲がりなりにも形になったのは素晴らしいです。 同じ班のメンバーの方々、ファシリテーターの松田先生のおかげでした。	
良いグループになったせいか、色々教えていただき、もっと勉強しなければいけないと気づかされました。 どういうことを考えて、研究を作り上げているのかを垣間見ることが出来、非常に良かったです。 事前学習があったから、よかったのかもしれませんが。	
自分では発想できないいろいろな意見が出て、とても勉強になりました。	
少人数グループだったので、分からない点や疑問点をその都度聞きながら進めることができ、大変勉強になりました。	
また、このような企画があれば参加したいです。	
グループで行うことで、自分だけでは気付かなかったことが気付けたのでよかった。	
いろいろな背景の方々と一緒にワークで来てよかった。	
ファシリテーターの先生と一緒に考えるスタンスで話につきあっていただいたことが、議論の自由度を上げたと思う。	
発表会に対する意見	
もう少しゆっくり他の班の発表やPECOを聴きたかったです。	
もう少し、時間をかけて、それぞれの内容を吟味できたらなおよかった。	
部屋を間違えて参加が遅れたため、他のグループの発表を聞くことができなかった。 そのため、何が良くて、何が悪かったのか理解することが難しかった。	
私達のグループでは出なかった意見や、そういう考え方もあるのかと納得させられる部分もあり、とても満足した。	
各班の回答を知ることができたが、かなり急ピッチで、後でどれがよかったか選ぶように言われてもほとんど覚えていられなかったのが残念だった。 セットアップしたもので、各班のRQを1枚のスライドにして選べたらもっとよかったと思う。	
各班のPICQのシートのコピーが欲しかった。 各班の良い点、悪い点を比較できたと思う。	
他のグループのリサーチクエスチョンをみることによって、様々な考え方があることがわかり、非常に勉強になりました。	
内容の詰めはともかく、作り上げていく課程をグループワークを通して体験できました。 コンテスト形式にしたのも盛り上がり良かったと思います。 各班独自の考え方が聞かれました。 ただ難をいえば、2日目の朝が早かったことぐらいでしょうか...	
他の方がどういうことを考えているのか少し知れたこと。	
優秀賞をとったもの以外は(自分の所属する班のものも含め)どんぐりの背比べだったと思います。 この辺の論評がもう少しあればよかったと思います。	
少々、時間が足りなかった感がありました。	
時間が無かったせいか、各グループの発表が十分でなく、せっかく作ったPECOが会場の皆さんに十分伝わらなかったため。	
グループで行ったことで達成感が味わえた。	
スライドだけでは、各グループがどんなPICQを立てたのか記憶できない。 で、自分たちのグループはどこがよくてどこが改善の必要があるのか、具体的にわからなかった。	
ロイトンまでの道を間違えて遅れました。 はじめのほうのグループだったので自分のとこが聞けなかったのが残念。	
良い評価をいただいた事はうれしかったが、上に書いたようにそれぞれの回答に対する、批評とアドバイスに時間をかけていただきたかったと思う。	

第 2 回日本プライマリ・ケア連合学会学術大会
 【ワークショップ】臨床研究デザイン道場「おけいこ編」
 【連動企画 IG】 臨床研究デザイン道場「仕上げ編」
 ～道場破り大歓迎～

2011 年 7 月 2 日(土) 9:00～11:00:ホール B 2 階 リージェントホール
 3 日(日) 8:00～9:00:第 1 会場 2F ハイネスホールA・B

■趣旨

「臨床研究は、結局統計解析なので苦手」と誤解されていませんか？良いリサーチ・クエスチョンを考えることは、臨床家のあなた以外にはできません。症例報告から一歩前へ。

さあやってみましょう！でも、その前に、分析的な研究の本質を理解する必要があります。データをとる前のデザインこそが重要です。臨床シナリオや抄録をケーススタディの題材として検討することを通じて、リサーチ・クエスチョンの構造化や臨床研究デザインの真髄を学んでいただくという企画です。

■アジェンダ・学習目標

1. ミニ・レクチャー: 日常の疑問をリサーチ・クエスチョンに構造化してみよう
・リサーチ・クエスチョンを構造化する際のガイド(PECO)の作り方や、臨床研究の基本的な考え方を解説する
2. グループディスカッション
・日常診療でありそうなシナリオの題材をリサーチ・クエスチョンに構造化してみよう
・より良いリサーチ・クエスチョンを作成するために、多職種で議論してみよう

■プログラム

司会進行:渡部	内容
初日 9:00	オープニング
9:05-9:35	ミニ・レクチャー (福原)
9:35-9:45	グループ実習の説明 シナリオ提示 アイスブレイク
9:45-10:45	グループワーク
10:45-11:00	初日の総括・アドバイス(野口)
15:00	グループワークプロダクト提出期限！！
2 日目 8:00	オープニング
8:05-8:45	グループワークプロダクト レビュー 質疑応答 優秀グループ選出の投票
8:50-9:00	優秀グループの表彰 総評

図 1 第 2 回日本プライマリ・ケア連合学会学術大会
 【ワークショップ】臨床研究デザイン道場概要

Ⅲ. 研究協力報告書

厚生労働科学研究費補助金（臨床研究基盤整備推進研究事業）
平成 23 年度 研究協力報告書

臨床研究に関するブレンデッド型遠隔学習プログラムの教育効果

研究協力者	角舘 直樹	京都大学大学院医学研究科	医療疫学分野	講師
研究協力者	次橋 幸男	京都大学大学院医学研究科	医療疫学分野	研究員
	山本 洋介	京都大学大学院医学研究科	医療疫学分野	講師
	三品 浩基	京都大学大学院医学研究科	医療疫学分野	
	横山 葉子	国立循環器病研究センター	日本学術振興会特別研究員	
	中村 文明	京都大学大学院医学研究科	医療疫学分野	
	福森 則男	京都大学大学院医学研究科	医療疫学分野	
	竹上 未紗	国立循環器病研究センター	予防医学・疫学情報部	研究員
	大野 慎也	京都大学大学院医学研究科	薬剤疫学分野	
	佐藤 恵子	京都大学大学院医学研究科	医療疫学分野	
	脇田 貴文	関西大学社会学部社会学科心理学専攻		助教
研究分担者	渡部 一宏	昭和薬科大学薬学部医療薬学教育研究センター		講師
研究分担者	山口 拓洋	東北大学大学院医学系研究科医学統計学分野		教授
研究代表者	福原 俊一	京都大学大学院医学研究科	医療疫学分野	教授

我々は、臨床現場で働く多忙な医療者が、臨床研究コア・コンピテンシーを学ぶための対面授業と遠隔学習システムを組み合わせたブレンデッド型遠隔学習プログラムを開発した。今回は、本プログラムの教育効果について報告する。

- 1) 本プログラム終了から4か月経過時点で、対象者の64%が臨床研究に取り組み始めていた。
- 2) 本プログラムは、臨床研究フェローシップの構築に必要な研究の仲間を増やすことにも寄与している可能性が示唆された。
- 3) 本プログラムは、受講者の臨床研究に対する自信（自己効力感）を高め、臨床研究実施を促進している可能性が示唆された。

A. 研究目的

医療者が日常診療の中でエビデンスを活用し、さらに自ら生み出すには、臨床研究のデザインから解析までの基礎知識（臨床研究コア・コンピテンシー）が重要である。特に、我が国

では臨床研究コア・コンピテンシーを理解する医療者が不足していることや、医療者からのニーズが高いことから、人材育成は重要な課題である。

近年、多忙な医療者が日常診療を継続しながら

臨床研究を系統的に学ぶための教育方略として遠隔学習システムの「e-learning」が注目されている。しかし、「e-learning」は対面授業と比較して完遂率が低いという短所があり、それを補うために遠隔学習と対面授業を組み合わせたブレンデッド型遠隔学習が注目されている。これまでに我々は臨床研究コア・コンピテンシーを習得し、自己効力感の向上を促進するためのブレンデッド型遠隔学習プログラムを開発した。自己効力感とは、「ある結果を生み出すために必要な行動をどの程度上手く行うことができるか」という自信であり、自己効力感を高めることで行動変容することが、医学・歯学・教育学・心理学領域で実証されている。本プログラムを実施した結果、多忙な医療者である受講者の多くがプログラムを完遂し、職種に関わらず臨床研究コア・コンピテンシーを習得することに成功した。

本研究では、知識の習得以外にもプログラム受講者が得た自己効力感や、受講後に臨床研究に取り組んでいるかなどの実践に関する教育効果を評価することを目的とした。

B. 研究方法

1. 対象者

本プログラムの受講者は、1) 医療機関に勤務している医療職である、2) 所属機関の推薦がある、3) 臨床研究の初学者であることの3条件を満たす者とした。参加募集は2009年7月1日から2009年8月15日までの期間に学術雑誌および専用のウェブサイトを通じて行った。

受講者176人の特性としては、男性が54%、平均年齢は35.0±7.9歳であった。職種としては医師が35%、薬剤師が24%、看護師が13%、理学療法士と栄養士／管理栄養士がともに5%であった。1週間の勤務時間については、

40時間以上と回答した者の割合は全体の83%であった。

2. 方法

1. ブレンデッド型遠隔学習プログラムの学習達成目標および教育内容

本プログラムでは、臨床研究のデザインおよび統計解析に関する基礎的知識の習得に主眼をおいた。講義は、平成21年9月から22年3月まで1か月に1回（土曜日）、全7日、合計23講義（1講義あたり60分）が提供された。

2. 本プログラムの教育方略

本プログラムでは以下の教育方略を用いた。

1) 同時ライブ配信方式での遠隔講義

ウェブを用いたTV会議システムを採用し、京都大学において行われた講義を全国7ヶ所のサテライト会場に同時ライブ配信を行った。

2) ウェブを用いた受講者管理システム

受講者管理システムを開発し、各受講生に対して個人用ウェブサイトを提供して受講状況や理解度チェックテストの結果を通知した。また、このサイト上において各受講生による講義への評価や質疑応答が行われた。2009年9月から2010年3月までの7ヶ月間で、計3,564件のアクセス数が確認された。さらに、受講者から得られた質問の数は計284件であった。

3) 理解度チェックテストの実施

各講義の最後に、出席の確認を兼ねた理解度チェックテストを実施した。内容は選択式で5つの問題で構成され、講義内容を確認することを目的とした。

3. 本プログラムの教育効果

3-1. 調査項目

本研究では以下の3つの観点から、本プログラムの教育効果を評価した。

1) 臨床研究の取り組みに対する効果

プログラムを受講後、自分が中心となって一番力を入れている臨床研究の進捗について、①リサーチ・クエスチョンを作った、②研究計画を完了した、③データ収集を完了した、④データ解析を完了した、⑤学会発表抄録を提出した、⑥学会発表をした、⑦論文を作成して、投稿した、⑧研究予定なしのいずれか一つに回答することとした。

2) 臨床研究フェローシップ構築に対する効果

臨床研究について議論できる仲間の人数を、プログラムの受講以前と受講4か月後について、該当する人数を記入した。

3) 臨床研究に対する自信（自己効力感）

現在の臨床研究実施に関する自信（自己効力感）の程度について、①リサーチ・クエスチョンを立てられる、②適切な臨床研究デザインを選択できる、③研究プロトコールを書ける、④データ測定信頼性と妥当性がわかる、⑤研究で生じるバイアスに対処できる、⑥研究で生じる倫理的問題に対処できる、⑦調査用の質問紙が作成できる、⑧研究結果を論文にできる、の8項目に回答してもらった。回答選択肢は、「出来る」から「まったく出来ない」までの5段階のLikert方式で評定した。

3-2. 調査方法

臨床研究に対する自信（自己効力感）は、プログラム受講前・受講直後・受講後4か月後の3回測定を行った。他の項目については、プログラム受講4か月後にのみ調査した。

4. 統計学的解析

臨床研究の取り組みに対する効果については

各回答項目を選択した者の割合を算出した。臨床研究フェローシップ構築に対する効果については、受講前と受講4か月後の臨床研究について議論できる仲間の人数の平均値の差を対応のあるt検定にて比較した。

臨床研究に対する自信（自己効力感）については受講前と受講直後、および受講前と受講4か月後の平均値の差を対応のあるt検定にて比較した。有意水準は、 $p < 0.05$ とし、データ解析には、STATA/IC, version 10.1 (STATA Corp LP., TX, USA)を用いた。

5. 倫理的配慮

本研究は京都大学大学院医学部医の倫理委員会の承認を得た後に実施された。本研究への参加同意に関しては、参加者の自由意志で行われることを書面にて通知し、同意を得た。

C. 研究結果

結果プログラム受講者176名のうち、プログラム受講4か月後の調査に回答したのは129名(73%)であった。また、臨床研究に対する自信（自己効力感）について、プログラム受講前・受講直後・受講後4か月後の全3回の調査すべてに回答したのは104名(59%)であった。

1. 臨床研究の取り組みに対する効果

プログラム受講4か月後の臨床研究の取り組みに対する変化については、「リサーチ・クエスチョンを作った」、「研究計画を完了した」、「データ収集を完了した」、「データ解析を完了した」、「学会発表抄録を提出した」、「学会発表をした」、「論文を作成して、投稿した」、「研究予定なし」と回答した人がそれぞれ30名(23%)、10名(8%)、13名(10%)、4名(3%)、17名(13%)、6名(5%)、3名(2%)、46名(36%)であった(図1)。

2. 臨床研究フェローシップ構築に対する効果

臨床研究について議論できる仲間の人数の平均値は、プログラムの受講前および受講4か月後の調査時点でそれぞれ、 2.6 ± 4.2 (SD)、 5.0 ± 6.1 (SD) 人であり、受講4か月後の方が統計学的に有意に大きかった ($p < 0.001$)。

3. 臨床研究に対する自信（自己効力感）

8項目すべてにおいてプログラム受講前と比較して、プログラム受講後および受講4か月後の臨床研究に対する自信（自己効力感）が統計学的に有意に高かった（表1）。

D. 考察

考察

本研究では、医療者を対象に、臨床研究コア・コンピテンシーの継続的な学習を目的としたブレンデッド型遠隔学習プログラムの教育効果を検討した。その結果、受講者の臨床研究の取り組みや臨床研究に対する自信（自己効力感）が高まり、実際に臨床研究を実施していることが示唆された。また、本プログラムを通じて、臨床研究について議論できる仲間が約2倍に増加し、臨床研究フェローシップが構築されたことが示された。

本研究結果では、プログラム受講により臨床研究に対する自信（自己効力感）が統計学的に有意な改善を認めただけでなく、その改善の度合（効果量）は1SDに匹敵する程の大きさであった。自己効力感を高める情報源としては①成功体験、②代理的経験、③言語的説得、④生理的・情動的状態がある。本プログラムでは、理解度チェックテストに合格することでスモールステップの成功体験を反復して積んでもらい、受講管理システムでの質問に対するフィードバック（言語的説得）や毎回の講義における

受講者間の交流（代理的経験）によっても、受講生が励まされたと思われる。これらの効果により、受講生の自己効力感が高まり、実際に自ら臨床研究を実施して周囲に臨床研究の仲間を増加させたものと考えられる。また、ブレンデッド型遠隔学習プログラムでは、従来のe-learningと異なり、他の受講者と同じ教室で授業を受けることから、フェローシップを構築しやすいと考えられた。

E. 結論

本研究の結果、医療者に対してブレンデッド型遠隔学習プログラムにて臨床研究教育を提供することで、周囲に研究を行う仲間が増え、さらに臨床研究実施が促進されることが示唆された。このようなプログラムに基づく教育システムが構築されることで我が国に強固な臨床研究の基盤が築かれることが期待される。

F. 研究発表

1. 論文発表

Tsugihashi Y, Kakudate N, Yokoyama Y, Yamamoto Y, Mishina H, Fukumori N, Nakamura F, Takegami M, Ohno S, Wakita T, Watanabe K, Yamaguchi T, Fukuhara S. A novel internet-based blended learning program providing core competency in clinical research. *Journal of Evaluation in Clinical Practice* 2011 (in press).

角舘 直樹, 次橋 幸男, 横山 葉子, 山本 洋介, 三品 浩基, 中村 文明, 福森 則男, 竹上 未紗, 大野 慎也, 佐藤 恵子, 脇田 貴文, 渡部 一宏, 山口 拓洋, 福原 俊一. 臨床研究に関するブレンデッド型遠隔学習プログラムの教育効果—医療者に対する臨床研究教育は自己効力感および臨床研究実施を促進する—. *医学教育*

2012 (in press) .

2. 学会発表

Tsugihashi Y, Kakudate N, Yokoyama Y, Yamamoto Y, Mishina H, Fukumori N, Nakamura F, Takegami M, Ohno S, Wakita T, Watanabe K, Yamaguchi T, Fukuhara S. A novel internet-based blended learning programme providing core competency in clinical research. An International Association for Medical Education in Europe, Vienna, Austria (August, 2011).

H. 知的財産権の出願・登録状況(予定を含む)

1. 特許取得
2. 実用新案登録
3. その他

特になし

図1. プログラム受講4か月後の臨床研究の取り組み

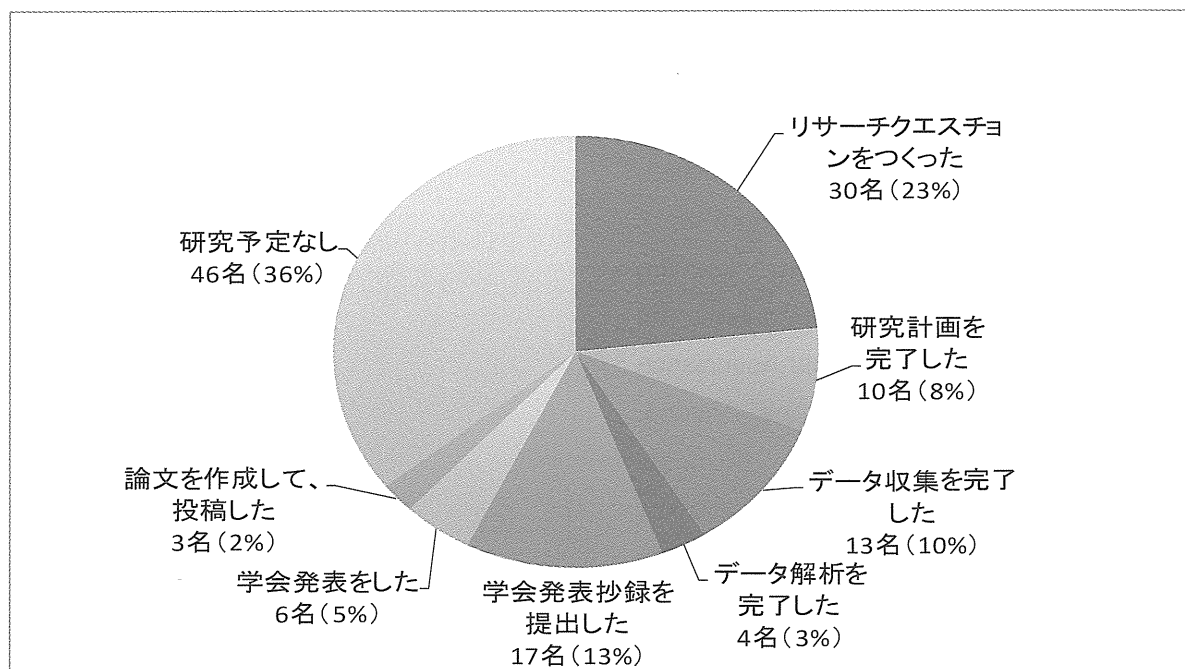


表1 臨床研究に関する自信（自己効力感）

	受講前	受講直後	P 値*	受講4か月後	P 値**
	平均値±標準偏差	平均値±標準偏差		平均値±標準偏差	
1. リサーチ・クエスチョンを立てられる	2.47±1.20	3.68±0.85	P <0.001	3.69±0.84	P <0.001
2. 適切な臨床研究デザインを選択できる	2.20±1.00	3.21±0.09	P <0.001	3.38±0.87	P <0.001
3. 研究プロトコールを書ける	2.12±1.05	3.02±0.10	P <0.001	3.16±0.94	P <0.001
4. データ測定の信頼性と妥当性がわかる	2.08±1.07	3.13±0.94	P <0.001	3.11±0.92	P <0.001
5. 研究で生じうるバイアスに対処できる	1.97±0.97	2.90±0.83	P <0.001	3.02±0.93	P <0.001
6. 研究で生じうる倫理的問題に対処できる	2.38±1.16	3.29±0.97	P <0.001	3.48±0.98	P <0.001
7. 調査用の質問紙が作成できる	2.20±1.05	3.17±0.97	P <0.001	3.35±0.89	P <0.001
8. 研究結果を論文にできる	2.34±1.29	2.83±1.06	P <0.001	2.99±1.07	P <0.001

*受講前と受講直後との平均値の差の検定(対応のあるt検定) **受講前と受講4か月後との平均値の差の検定(対応のあるt検定)

n=104

各項目は、「出来る：5点」、「少し出来る：4点」、「どちらともいえない：3点」、「ほとんど出来ない：2点」、「まったく出来ない：1点」の5段階で評定した。

臨床研究コア・コンピテンシーテストの検証

研究協力者 脇田 貴文 関西大学 准教授

研究協力者 山本 洋介 京都大学 医療疫学分野 講師

研究要旨

本研究課題において、開発した臨床研究コア・コンピテンシーテストに関して、テスト理論に基づく心理計量学的な検討を行った。具体的には個々のテスト問題に関して検討を行った後に、テストとしての性能を評価した。その結果、当該テストは信頼性の観点、因子的妥当性の観点から妥当なものであると判断された。

A. 研究目的

臨床研究を行う際に必要な知識を習得するためにさまざまな取り組みを行ってきたが、それらの取り組みを評価する際に2つの視点が存在する。1つ目は、学習者個人がどの程度知識を習得できたか、2つ目は、その学習プログラムがどの程度の効果を上げることができるかである。つまり、前者は個人、後者はプログラムそのものであり、これらは表裏一体のものであると言える。このように、臨床研究教育に対する評価を行う際の重要な鍵となるものが学習者に対する試験（テスト）である。臨床研究に対する確立した評価方法が定まっていない現状を鑑みると、まず、臨床研究を評価するためのテスト自体が、信頼性妥当性の観点から質の高いものかどうかを検証した上で、このテストを用いて対象者への臨床教育の効果を明らかにする、という手順が必要である。すなわち、臨床研究評価のためには、臨床研究教育の効果を測定する質の高いテストを作成することが何よりも重要であることは言うまでもない。

なお、テストを開発する際には、さまざまな

観点から検討を加える必要がある。まず作問段階では、何を測定したいのか（測定概念）を決定することが重要である。そして測定概念に基づき、問題（項目）を作成し、データを収集する。その後、項目、テストに関して検討を行い、さまざまな修正を加えより適切な測定をするためにテストを洗練していく必要がある。

本研究課題では、2009年度にwebを用いた臨床研究コア・コンピテンシーテストを開発した（山本・脇田・角館・福原, 2010）。本研究では、2009年度に引き続き、2010年度にもデータを収集し、臨床研究コア・コンピテンシーテストに関して、心理計量学的な検討を加え、より良いテストの開発につなげることを目的とする。

B. 研究方法

a) 対象者および実施時期

全国3会場および受験者自宅において、計168名に対して実施した。実施時期は2010年10月～2011年8月であった。

b) テスト内容

2009年度に開発した臨床研究コア・コンピテンシーテストを用いた。このテストは、分野「臨床研究デザイン」(以下、「デザイン(分野)」)、分野「統計解析」(以下、「統計(分野)」)の2つの分野から構成され、それぞれデザイン分野28問、統計分野22問の計50問であった。

c) 分析方法

以下の分析を行った。なお、紙面の都合上結果において省略したものがある。

【分析A】[古典的テスト理論に基づく分析]

- ・正答率(通過率)およびその度数分布
- ・点双列相関係数
- ・選択肢別選択率
- ・GP分析(成績群別の選択肢別選択率の比較)
- ・因子分析
- ・信頼性係数の推定
- ・正答数の度数分布
- ・正答数の平均、標準偏差

【分析B】[Item Response Theory(以下、)に基づく分析]

- ・識別力パラメタ、困難度パラメタ
- ・項目特性曲線
- ・テスト情報曲線

3.2. 使用ツール

分析を行うに当たり、目的に合わせて以下のソフトウェアを使用した。

- ・Microsoft® Office Excel 2003 (Microsoft社)
- ・PASW Statistics® 17.0 (IBM SPSS社)
- ・Sobi Version 1.0.0 (2値型IRT計算プログラム、開発協力:熊谷 龍一准教授/東北大学大学院教育学研究科)

(倫理面への配慮)

一連の本研究における、参加者の理解度や満足度を測定する過程は、京都大学医の倫理委員会に提出、承認済みである。また、webを用いたCBT(Computer Based Test)であるため、同意に関わる事項は画面にて表示し、「同意する」にチェックを付けたものを対象にテストを実施した。

C. 研究結果

i. 正答率(通過率)および点双列相関係数

各設問の正答率および点双列相関係数を表1に示した。全50問の設問別正答率の平均は0.675、標準偏差は0.186であり、点双列相関係数が0.20未満のものは4項目であった。特に設問ID2007は0.07と極端に低い値であった。除外後の46項目における点双列相関係数を併せて表1に示した。削除後においては点双列相関係数が0.20未満のものはなかった。

ii. 因子分析および信頼性係数の推定

削除後の46項目に関して、因子分析(主因子法)を行なった。固有値の減衰状況は14.78、3.07、2.44、…であった。複数因子の可能性も示唆されたが、2因子解を求めたところ解釈可能な因子ではなく、また因子間相関は0.599であったため、1次元であると判断した。

iii. IRTに基づく分析

各項目の項目パラメタを推定し、表2に示した。また、識別力および困難度をプロットした散布図を図1に示した。困難度は、-4.19から1.89に分布していた。また、識別力は0.22から

2.00 に分布していた。

図2にテスト情報曲線を示した。潜在特性連続体上で、テスト情報量が 9 を越えるのは -2.40 から 0.10 であった。

D. 考察

i. 正答率(通過率)および点双列相関係数

正答率の観点では、当該受験者集団においては妥当な困難度であったと考えられる。点双列相関係数に関しては、低いものがあったが、それらを削除することで、一定の値が得られた。なぜ、削除対象となった項目が、他の項目と異なったのかについては内容的な側面からの検討を加え、今後の問題作成にいかすべきである。

ii. 因子分析および信頼性係数の推定

因子分析の結果から、1 因子構造であることが示されたが、その理由は必ずしも積極的なものではない。できる限り項目を残すという方針で分析を行ったが、あと数項目削除を行えば、より1因子構造になるだろう。

信頼性の観点では、テストとして十分高い信頼性を示していると考えられる。

iii. IRT に基づく分析

各項目のパラメタから(図 1)、統計分野の識別力が、デザイン分野に比して低くなっている。この点が上記の 1 因子構造を弱めている原因となっている可能性がある。統計分野の項目は、臨床研究を行う上で必要なものではあるが、それ以外の場面でも用いられるものである。つまり、臨床研究に特化されていないという点がこのような結果をもたらしている可能性がある。

また、テスト情報量をみると、-2.4~0.1 の範

囲で、かなり精度の高い測定を行うことができるテストとなっている。臨床研究コア・コンピテンシーテストが、受験者の相対的な比較を目的にしているわけではなく、学習の到達度を測定していることを考慮すると妥当なものであると思われる。しかし、潜在特性値が 0.5 から 1 程度(偏差値では 55 から 60)程度の受験者に対する測定精度を上げることが必要であろう。

E. 結論

臨床研究コア・コンピテンシーテストは、心理計量学の観点から、十分な信頼性をもち、実用的な水準に達しているものと考えられる。一方、より適切な測定、評価を行うために、項目の入れ替え等を行う必要があるだろう。

また、これらの結果を踏まえて、項目の内容面にまで踏み込んだ検討を続けていく必要があると考えられる。

F. 研究発表

1. 論文発表

なし

2. 学会発表

(発表誌名巻号・頁・発行年等も記入)

なし

H. 知的財産権の出願・登録状況(予定を含む)

1. 特許取得

2. 実用新案登録

3. その他

なし

表 1 各項目の正答率と点双列相関係数(除外前・除外後)

分野	設問ID	正答率	点双列相関係数		分野	設問ID	正答率	点双列相関係数	
			除外前	除外後				除外前	除外後
統	2001	0.77	0.41	0.41	デ	2026	0.80	0.64	0.65
統	2002	0.86	0.17	—	デ	2027	0.80	0.51	0.51
統	2003	0.67	0.24	0.23	デ	2028	0.86	0.38	0.39
統	2004	0.83	0.18	—	統	2029	0.93	0.39	0.39
統	2005	0.80	0.31	0.31	デ	2030	0.97	0.46	0.47
統	2006	0.89	0.28	0.28	デ	2031	0.70	0.35	0.36
デ	2007	0.55	0.07	—	デ	2032	0.48	0.50	0.50
統	2008	0.46	0.32	0.33	デ	2033	0.91	0.34	0.35
デ	2009	0.38	0.19	—	デ	2034	0.45	0.43	0.43
デ	2010	0.81	0.50	0.51	デ	2035	0.57	0.53	0.54
デ	2011	0.89	0.22	0.22	デ	2036	0.85	0.44	0.44
統	2012	0.65	0.35	0.35	デ	2037	0.73	0.50	0.49
デ	2013	0.32	0.25	0.25	デ	2038	0.48	0.31	0.31
統	2014	0.68	0.38	0.38	デ	2039	0.48	0.43	0.44
統	2015	0.56	0.42	0.41	統	2040	0.41	0.35	0.34
統	2016	0.72	0.41	0.42	デ	2041	0.20	0.27	0.27
デ	2017	0.44	0.51	0.50	デ	2042	0.94	0.40	0.40
統	2018	0.33	0.31	0.33	デ	2043	0.80	0.40	0.40
デ	2019	0.79	0.39	0.39	デ	2044	0.59	0.49	0.50
デ	2020	0.74	0.30	0.32	統	2045	0.83	0.46	0.47
デ	2021	0.67	0.48	0.48	デ	2046	0.77	0.42	0.43
統	2022	0.55	0.48	0.48	統	2047	0.72	0.36	0.36
統	2023	0.47	0.42	0.43	統	2048	0.79	0.47	0.48
統	2024	0.76	0.52	0.52	デ	2049	0.86	0.49	0.50
統	2025	0.70	0.49	0.48	統	2050	0.57	0.27	0.27