平成19年度国際共同研究

医師の客観的臨床教育能力評価に関する研究
–OSTE：Objective Structured Teaching Evaluationの試み–

長崎大学病院 医師育成キャリア支援室　准教授

浜田　久之

【スライド -1】

このような機会を与えていただき、有り難うございます。

長崎はやはりかなり医師不足でして、非常に大変な状況になっています。

私は長崎医療センターから長崎大学へ移りましたが、長崎大学は西洋医学の発祥の地でありますし、また、長崎医療センターではスーパーローテイトシステムを 40 年前くらいに始めました。歴史と伝統のある長崎から医学教育の情報発信をできないかということが発端で、研究を始めております。

【スライド -2】

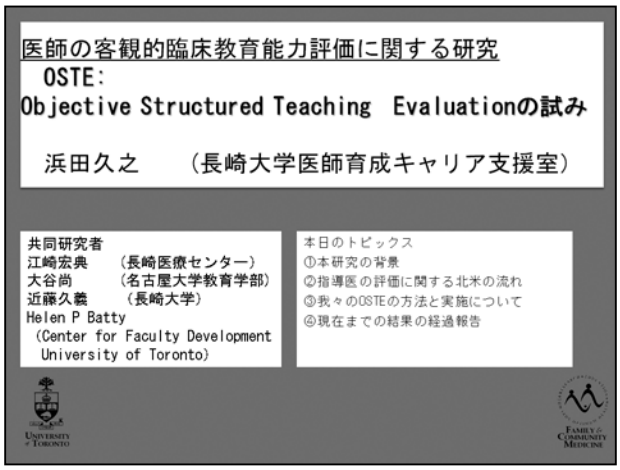
私の留学先であるカナダのトロント大学と一緒に研究をしてきました。

良い臨床の先生が集まるところに研修医も集まるし、学生も集まるということが、元々の発想であります。

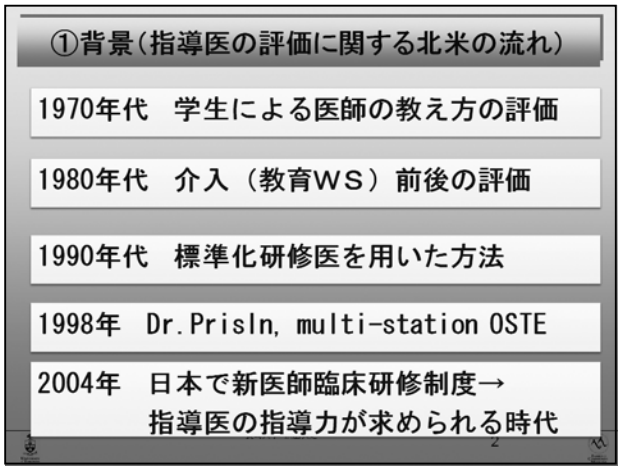
いろいろ向こうで私も経験したのですが、かなり進んでおりまして、トロント大学では学生や研修医から教え方の評価を受けて、それが昇進や給料の査定になるというところまでいっていますので、驚きました。日本でも、2004 年に臨床研修医制度が始まって、我々が研修医を評価しなければならないことになったのですが、我々は安全な立場にいて、研修医から評価されることはほとんどありません。しかし、今後はおそらく我々も評価されるのではないかということで、トライをしてみました。

日本では文献的にもトライアル的にもあまりやっていないので、

スライド-1

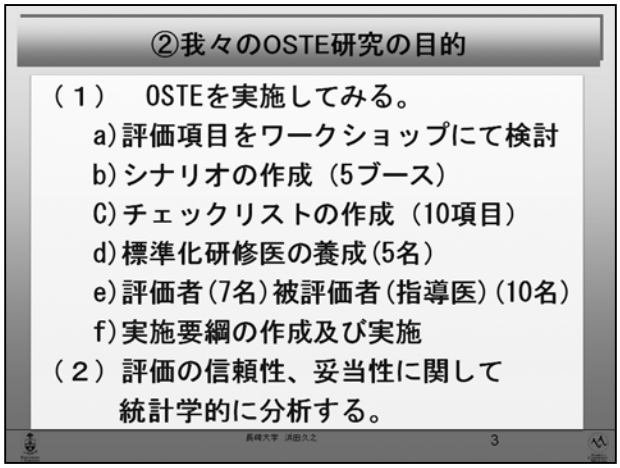


スライド-2



とりあえずやってみようということが一番大きく、あと、どういう反応が出るかということ、そして統計学的にそれを分析してみようということがありました。

スライド-3



【スライド -3】

我々がやったことは、何をもって良い先生とするかという項目を作って、それに対するシナリオを作り、指導医の先生たちがそのシナリオに沿った教え方をして、それをチェックするということです。そのためには研修医の役者を養成しなければいけませんので、その役者さんを養成する。元々の研修医の先生たちを 6 時間くらい、2 時間 ×3 回で養成をしました。そして評価者を選ぶ。もちろんそれに加わってくださる指導医の先生方をリクルートするということが、一年間の仕事でありました。

【スライド -4】

スライドは、細かいですけれども、評価項目です。これが一番大変なところでした。どうやって評価項目を作るかということについては、海外の文献を十分に参考にしたということと、我々の教育研究所でカナダの先生たちを呼んで数回ワークショップを開いたりとか、また、過去に指導医講習会をたくさんやっており、そこに " 良い指導医・悪い指導医 " というセッションがあるのですが、そこからたくさん抜き出しました。これは質的研究で抜き出そうと思ったのですが、なかなかうまくいかなくて、皆で合意して作りました。

指導評価項目としては、「研修医が学習しやすい雰囲気を作る」、そして「研修医が指導するための教育的な知識を持っている」。あと、「研修医を指導するためのフィードバック技術を持っている」、そして、「医学的知識・技術を教えることができる」、また、「ロールモデルとなるような態度を示す努力をしている」ということにしました。

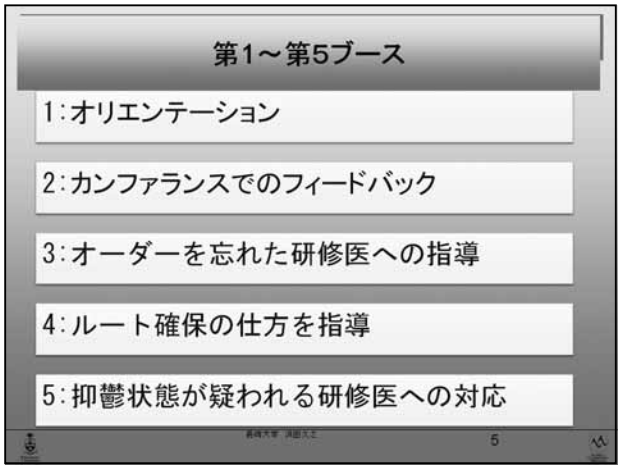
そして、これに対して質問項目ブースを作って質問項目をたくさん作りました。CCF は Cleveland Clinic 財団という向こうの財団が作った評価項目で、SFDP はスタンフォード大学の評価項目なのですが、どれに一致するかということを検討して質問をどんどん作っていきました。できるだけまんべんなく網羅するようにということで、評価項目を作っていきました。

スライド-4

【スライド -5】

どのようなブースを作ったかといいますと、第 1 ブースはどの研修も指導医も経験するオリエンテーション。また、カンファレンスでプレゼンテーションしますが、そこでのフィードバック。研修医の先生は忘れることがよくあるのですが、その忘れたときに指導医の先生がどうやって怒るか。点滴をするときにどうやって指導するか。あと、研修医の 50％くらいが鬱状態になると言われており、そういうことがよく起こるので、それにどうやって対応するか。それらのブースを作りました。

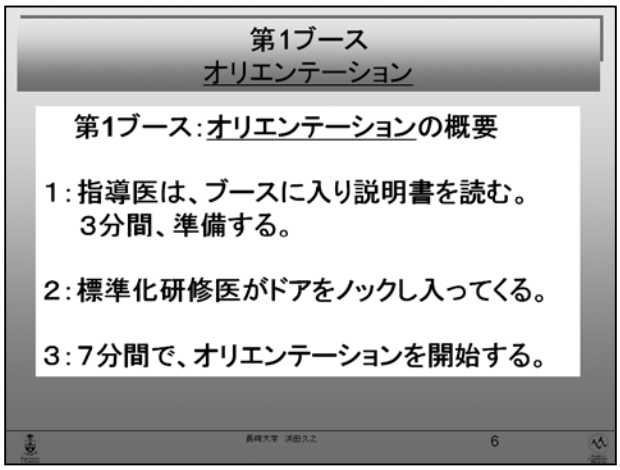
スライド-5



【スライド -6】

第 1 ブースはオリエンテーションですが、ブースの中に指導医の先生が入ってきて、どういうことが起こるかという説明書きを読んで、3 分間準備する。その後に標準化研修医が入ってきていろいろな質問をする。例えば、「1 週間後に友達の結婚式があるのですが、出席していいですか」とか「自分は学会で発表したい」とか、いくつかの質問を持っており、それを質問をする。指導医は 7 分間、パフォーマンスします。

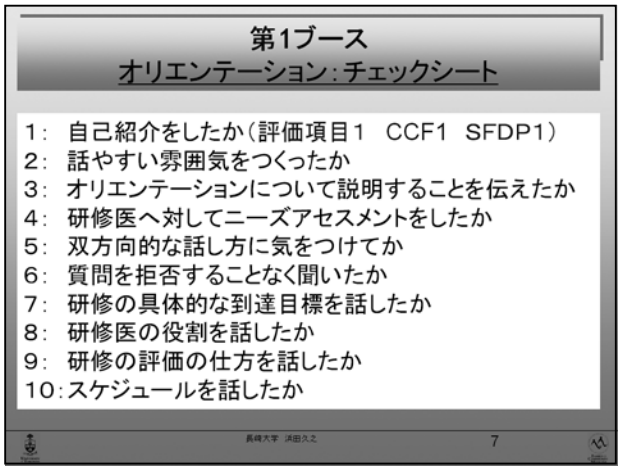
スライド-6



【スライド -7】

チェックリストは、「研修医に対して自己紹介をしたか」とか「話しやすい雰囲気をつくったか」とかで、10 項目あります。これが評価項目のどれに一致しているかは、スライド 4 で示しました。

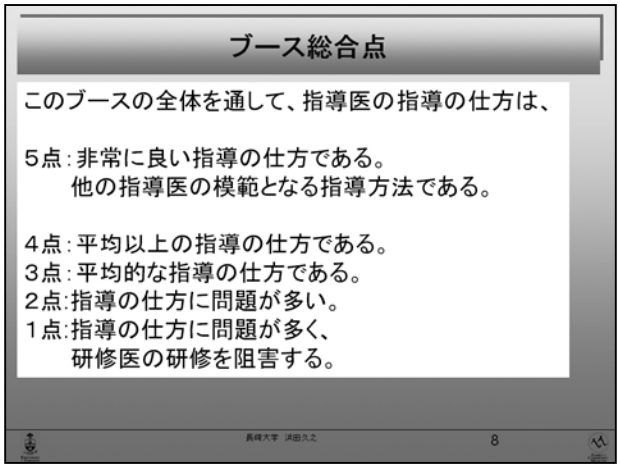
スライド-7



【スライド -8】

そのチェックリストと全体のパフォーマンスを総括評価しました。「非常に良かった」から「非常

スライド-8



スライド-9

被評価者（指導医）背景

指導医番号	性	科	年齢	臨床年数	指導経験年	年間指導研修医数	講習会受講	研修医指導が好き	研修医指導に自信がある	研修医指導で悩んだ経験は	筆記テスト結果(10点満点)
1	男	外科	40代	20年>	10−15年	10−15人	○	○	○	よくある	9
2	男	病理	50代	20年>	5−10年	1−5	○	○	○	時々ある	6
3	男	内科	30代	10−15年	5年以下	15−20	×	○	○	まったくない	5
4	男	内科	30代	10−15年	5−10年	20人以上	○	○	○	よくある	4
5	男	内科	20代	5年<	5年以下	15−20人	×	◎	○	時々ある	5
6	男	小児	30代	5−10年	5−10年	5−10人	○	○	△	時々ある	3
7	男	内科	30代	5−10年	5年以下	1−5人	○	△	△	時々ある	8
8	男	内科	40代	15−20年	5年以下	0人	×	○	×	よくある	10
9	男	内科	40代	10−15年	5年以下	1−5人	○	○	△	よくある	9
10	男	救命	40代	20年>	15−20年	20人以上	○	◎	○	時々ある	8

に悪かった」で5段階評価をしております。

【スライド-9】

どういう人がこれに加わってくれたかということですが、100名ぐらいの人にやってくれと頼んだのですが、結果的に10名の人が参加しました。すべて男の人で、いろいろな科の人がおります。20代から50代まで。臨床経験年数とか指導経験年数、また、指導医講習会を受けているか。あと、事前のアンケートをしました。この表の「筆記テスト」というのは、指導医講習会のプレテストで10点満点で教育的知識があるかどうかということも計りました。

スライド-10

評価者背景

	年齢	M/F	職業	研修医指導歴	年間指導研修医数	OSCE評価者研修受講歴	指導医講習会受講歴
評価者1	50代	M	医師	20年以上	20名	あり	あり
2	40代	F	看護師	0	0	なし	あり
3	40代	M	医師	10-20年	30名	なし	あり
4	30代	F	医師	5-10年	30名	あり	あり
5	50代	M	教員	0	0	なし	なし
6	40代	M	医師	10-20年	10名	あり	なし
7	20代	F	事務職	0	0	なし	なし

10

【スライド-10】

どういう人が評価をしたかということですが、これは医者もいましたし、教育専任看護師も臨床研修担当事務職もいて、ばらばらです。学生のOSCEの評価者の講習会を受けている人も何名かいました。そういうことで、あとで統計解析をしております。

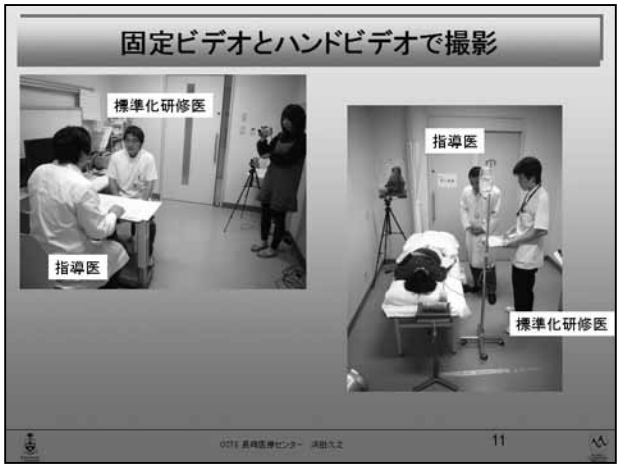
【スライド-11】

2台のデビオカメラで撮影し、標準化研修医、指導医、患者さんもいるという形で、このような5ブースを作りました。

【スライド-12】

1ヶ月後に、7人の評価者がDVDでパフォーマンスを見るということにしていました。

スライド−11

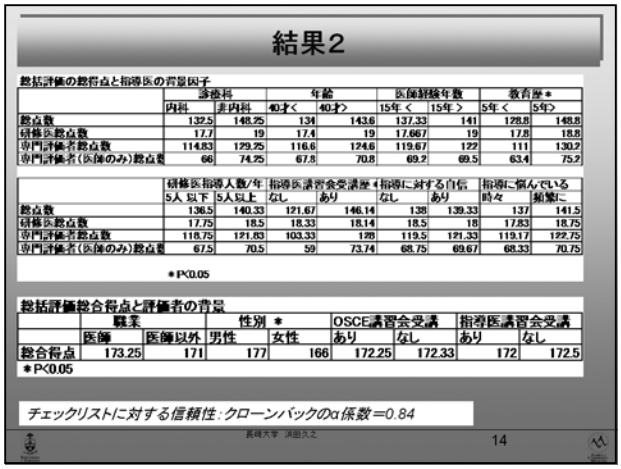


スライド−12

スライド−13

評価項目	評価者	平均 総合得点	統計学的結果
シナリオの現実性について（４点満点）	指導医	3.34 18.7	P>0.5 三者間で有意差なし
	専門評価者	3.14 15.7	
	標準化研修医	3.1	
研修医の演技に対する評価（４点満点）	指導医	3.12 15.8	P>0.5 三者間で有意差なし
	専門評価者	3.00 15.0	
	標準化研修医	3.2	

スライド−14



	診療科		年齢		医師経験年数		教育歴＊	
	内科	非内科	40才<	40才>	15年<	15年>	5年<	5年>
総点数	132.5	148.75	134	143.6	137.33	141	128.8	148.8
研修医総点数	17.7	19	17.4	19	17.667	19	17.8	18.8
専門評価者総点数	114.83	129.25	116.6	124.6	119.67	122	111	130.2
専門評価者（医師のみ）総点数	66	74.25	67.8	70.8	69.2	69.5	63.4	75.2

	研修医指導人数/年		指導医講習会受講歴＊		指導に対する自信		指導に悩んでいる	
	5人 以下	5人以上	なし	あり	なし	あり	時々	頻繁に
総点数	136.5	140.33	121.67	146.14	138	139.33	137	141.5
研修医総点数	17.75	18.5	18.33	18.14	18.5	18	17.83	18.75
専門評価者総点数	118.75	121.83	103.33	128	119.5	121.33	119.17	122.75
専門評価者（医師のみ）総点数	67.5	70.5	59	73.74	68.75	69.67	68.33	70.75

＊P<0.05

	職業		性別 ＊		OSCE講習会受講		指導医講習会受講	
	医師	医師以外	男性	女性	あり	なし	あり	なし
総合得点	173.25	171	177	166	172.25	172.33	172	172.5

＊P<0.05

【スライド -13】

結果です。

「研修シナリオの現実性」、「研修医の演技」は４点満点で３点台の半ばでまずまずということで、これは他の文献でもだいたいこのようなものでした。

【スライド -14】

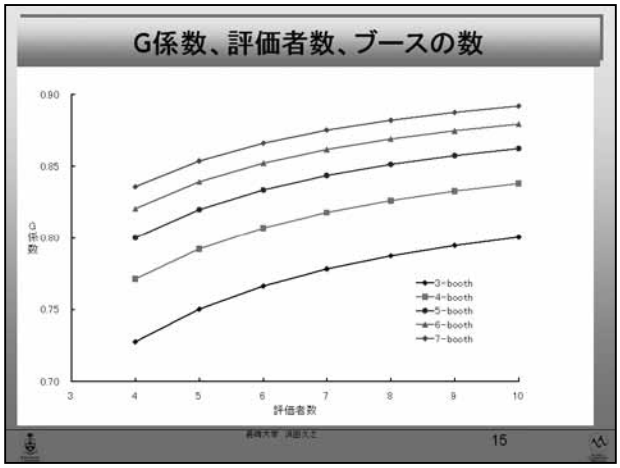
「総括評価の総点数と指導医の背景因子」では、内科系の先生が若干点数が良くて、教育歴が長い先生、指導医講習会を受けた先生の方が点数が良かった。

そして「総合評価と評価者の背景」ですが、女性の評価者が厳しく評価をしたということが、統計学的に証明されました。

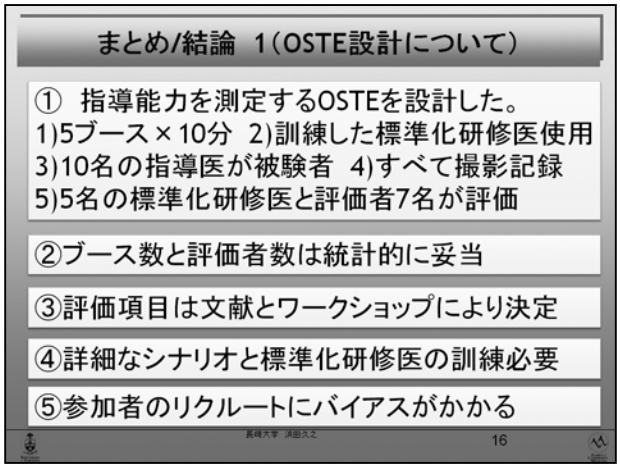
そしてチェックリストに対するα係数を算出したところ、0.84 でした。0.8 以上がまずまず信頼性があるということです。

又、チェックリストの総点数と総括評価の総得点も強い相関があり、OSTE の妥当性も証明されたと思います。

スライド-15



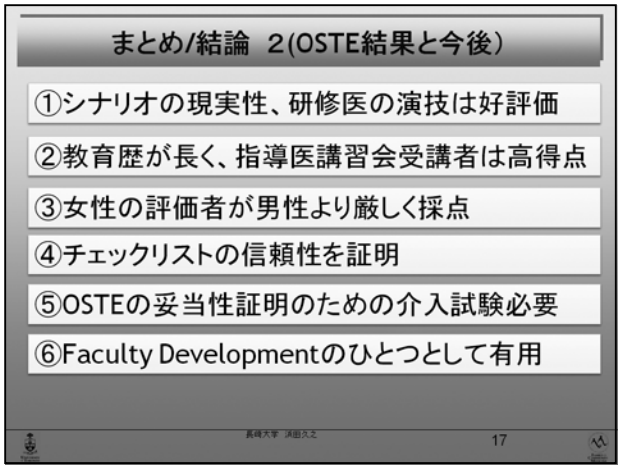
スライド-16



【スライド -15】

評価の人数とブースの数を統計学的に検討しましたが、評価者をたくさん増やすよりもブースの数を増やしていった方が、信頼性が上がるということが分かりました。ただ、なかなかブースを増やすのは大変ですので、我々は７人と５ブースということでした。

スライド-17



【スライド -16】

まとめです。

OSTE をこのようにやっていったということと、準備が１年間かかって大変であったということ。詳細なシナリオと訓練がかなり必要だということ。参加した人おそらく教育熱心な参加者ばかりだったのでバイアスがかなりかかっているだろう。

【スライド -17】

いろいろな結果が出ましたけれども、統計学的にチェックリストの信頼性は証明されて、今後は介入試験をしたりとか、あるいは何とか指導医講習会の一つのセクションとして利用するという計画を、今立てております。

質疑応答

会場： ここで取り上げられたのは、医学の学生たちがやられている OSCE（オスキー）の卒業版ということだと思いますが、人が人を評価するというのは大変難しい

作業で、私もその方法論を随分工夫してやっております。

ご発表は統計的に処理されているのですが、先生が実際にやられた統計的でない印象として、ここで「良かった」という評価の指導医は、先生が見て本当に良かったと思われますか？

浜田： 点数が高い先生は、やはり本当に研修医に人気がある先生でした。そういう先生はそういう技術を持っているのか、あるいは訓練してそうなったのか分かりませんけれども、点数が高い先生はそうでありました。ただ、低い先生はどうかというと、そこは何とも言えないところです。

会場： そこが問題ということですか。

浜田： はい、問題だと思います。しかし、チェックリストの総点（データーは示してない）と総括評価の総得点は、統計学的に相関しており、誰が評価しても同じような順位がついているという結果になっているので、信頼性は高いと思います。

会場： そこをきちっと判定してくれるようなチェックリストがあると大変有り難いなと思います。指導医講習会やあちこちで使わせてもらいたいと思って聞いていたのですが、まだ使うにはどうかなということですね。

浜田： そうですね。まだまだ何度かやっていかないといけません。ただ、「この先生がいい」とか「悪い」とかは、かなり研修医の判断で分かれるところがあります。非常に点数が低い先生もいらっしゃったのですが、ある意味、その人たちはぶっきらぼうであるとか。要するにこのテストは、親身になってやる人は点数が高く出るのです。だから、そういう指導を嫌う研修医もいるので、何とも言えないところもあります。しかし、指導方法の訓練としては十分に使えると思います。このチェックリストを使用すれば、上手な指導の仕方をすると、誰が評価しても‘上手’と評価できると思います。

会場： 是非、大学の教授でやってみていただきたいと思います。その結果は、来年にでも発表してもらったら大変有り難いと思います。（笑）

浜田： かなり大変だと思いますけれども、やってみたいと思います。

会場： 今回は標準化研修医ということで、実際には研修医の方を 6 時間位指導される、とお聞きしましたが、模擬患者さんというものをこういった医学教育で使われたりしていると思うのです。例えば、一般の方に標準化研修医になってもらうような指導をすると、毎年使えるのかなと思ったのですが、それは難しそうでしょうか？

浜田：　海外の文献でも、ほとんどが学生とかシニアレジデントを標準化研修医にしています。何故かというと、やはりかなり医学的知識がないと難しいということです。例えば、「採血のオーダーを忘れた」という理由の述べ方や、例えば鬱状態のなり方とか、そういうのはやはり研修医を経験するとリアルなのです。ものすごく演技が上手で、皆さん褒めていました。

座長：　今、いろいろな職種で現場での臨床指導あるいは研修が求められているのですが、医師ではない職種でも活用できるだろうということが何かありましたら、お教えください。

浜田：　一つは、看護師さんに評価者に入っていただいたのですが、やはり看護師さんのプリセプター教育ですね。私も看護師さんの FD をやっているのですが、そこらあたりが一番よいのではないか。あと、中堅の看護師や師長さんとかに。まあ、どの業種でもやれるのではないかなと思っています。