

科学研究費助成事業 研究成果報告書

平成 28 年 9 月 14 日現在

機関番号：11301

研究種目：基盤研究(C) (一般)

研究期間：2013～2015

課題番号：25460609

研究課題名(和文) 医学教育アウトカム評価法としてのカルテピアレビューシステムの確立

研究課題名(英文) Establishment of a peer review system using patient records for outcome evaluation of medical education

研究代表者

亀岡 淳一 (Kameoka, Junichi)

東北大学・医学(系)研究科(研究院)・准教授

研究者番号：30261621

交付決定額(研究期間全体)：(直接経費) 3,700,000円

研究成果の概要(和文)：信頼性の確立している「診療録ピアレビューシステム」の基準連関妥当性を検討した。協力病院(関東・中部地方の3病院)の内科外来を新患受診し、3年目研修医13名が診療した患者65名(各研修医5名)の外来診療録を、評価者5名が病院を訪れて評価し、各病院の指導医が別に行った研修医評価との相関係数を解析した。「医療面接」($r=0.509$)、「臨床推論」($r=0.585$)、「概略評価」($r=0.306$)等の相関は高く、「身体診察」($r=0.132$)、「患者・家族への態度」($r=-0.089$)の相関は低かった。主要な項目で基準連関妥当性が認められたが、診療録からは評価が困難な項目も明らかとなった。

研究成果の概要(英文)：We examined criterion validity of our peer review system of medical records as an outcome evaluation of medical education. We randomly selected 65 patients (five per senior resident) who were seen by 13 senior residents in three hospitals in Kanto and Chubu regions. Five reviewers visited the hospitals and evaluated outpatient medical records based on an evaluation sheet, of which reliability has been established. Independently, program directors of the 13 senior residents evaluated their clinical performances. The correlation coefficients of “history taking ($r=0.509$)”, “clinical reasoning ($r=0.585$)” and “overall assessment ($r=0.306$)” were high, while those of “physical examination ($r=0.132$)” and “attitude towards patients and family members ($r=-0.089$)” were low. We have established criterion validity of major items in our peer review system, but we also clarified that some items were difficult to evaluate by reading medical records.

研究分野：医学教育学

キーワード：アウトカム評価 診療録 ピアレビュー 基準連関妥当性

1. 研究開始当初の背景

教育の評価には、インプット評価(どのような教育が施されているか)、アウトプット評価(終了時にどのような能力が獲得されたか)、アウトカム評価(長期的にどのような人材を生み出したか)の3種類がある。日本の医学教育は、欧米の医学教育を土台にさまざまな改革が進められつつあり、そのインプット評価・アウトプット評価は検討されてきているが、アウトカム評価はほとんどなされていない。

例えば、平成16年度に厚生労働省により導入された新しい医師臨床研修制度のアウトプット評価(大学病院研修医3640名のアンケートによると、clinical competencyは著しく改善されたとの報告)はなされているが(Nomura K, et al. Med Teach 30:161, 2008)、アウトカム評価からはなされていない。

これまで、アウトカム評価は主に北米で検討されてきたが、explicitな基準(具体的な細かい項目に従った評価)では信頼性は上がるが妥当性が低く、implicitな基準(行間を読むような評価)では妥当性は高いが信頼性が落ちるというdilemmaが指摘されていた(Goldman RL. JAMA 267:958, 1992)。一方日本国内でも、いくつか試みがなされているが、例えば、外傷死亡患者1757例の質評価としての同僚評価peer reviewは、死亡例という特殊例に絞って解析されており(大友康裕他、日本外傷学会雑誌16:319, 2002)、遡及的診療録レビューの手法の妥当性の検証は、医療事故報告書との照合を通しての研究であり(Kobayashi M, et al. J Evaluation Clin Practice 14:126, 2008)、いずれも特殊事例に偏った評価であった。

そこで、我々は、できる限り患者の入口から

のrandom selectionを取り入れ、implicitな要素も加えた総合的な診療評価方法の開発を目標とし、2010年度に基盤研究(C)科学研究費(22590448)を獲得し、以下の3段階の計画を作成した。

第1段階として、信頼性の検討を目的にretrospective studyを行う。

第2段階として、基準連関妥当性の検討を目的に、多角的方面(患者、同僚、comedical staff等)からの相関を検討するprospective studyを行う。

第3段階として、第1・2段階で信頼性と妥当性の確認されたシステムを用いて、医学教育に関わる種々の因子(初期研修義務化の前後の比較、OSCE導入前後の比較等)の診療への影響を検討する。

まず、2013~2015年度に第1段階の信頼性のretrospective studyを行った。過去5年間に研究協力病院(東北地方の4病院)の内外来を新患受診し、後期研修医(卒後3~10年目)が診療し、最終的に入院となった患者112名(疾患名を問わない)の外来診療録を、評価者(東北地方以外の総合診療医4名)が、病院を別々に訪れて、評価表(記載の有無:3段階評価14項目、診療内容:5段階評価15項目)に基づいて評価した。結果は、評価者の相の級内相関係数の単一測定値は0.733、平均測定値は0.917であり、極めて良好な信頼性が得られた(Kameoka J, et al. Tohoku J Exp Med. 233:189-95, 2014)。

そこで、次に、第2段階の基準連関妥当性の検討が求められる。

2. 研究の目的

まず、2012年度に、第2段階の基準連関妥当性の検討目的で、東北大学および東北地方の協力病院を対象にprospective studyを計画

し、東北大学大学院医学系研究科倫理委員会に申請し、承認が得られた(2012-1-111)。しかし、この研究は、多角的方面(患者、同僚、comedical staff等)からの参加を要し、多大なる労力を有するため、関連病院からの協力を得られず、このままの方法では断念せざるを得なかった。

そこで、外的基準を「医師(指導医)による評価」に絞り、EPOC評価も加え、対象を(初期研修を終えたばかりの)3年目後期研修医に限定し、かつ、周囲の労力の比較的少ない retrospective study に変更して、基準連関妥当性の検討を行うことにした。

3. 研究の方法

【評価表】

(1)カルテピアレビューの評価表は、基盤研究(C)科学研究費(22590448)で使用した評価表を用いる。

(2)「指導医による研修医評価表」は、コアメンバー委員会で検討し、以下のように作成した。

- P1 病歴聴取
- P2 身体診察
- P3 臨床推論
- P4 治療
- P5 医療技術
- P6 カルテ記載
- P7 患者・家族への態度
- P8 他医・comedical staffとの連携
- P9 患者のアウトカム
- P10 全体の概略評価

以上を、(5:優れている、4:やや優れている、3:ふつう、2:やや劣っている、1:劣っている)の5件法で評価する。

【対象】

(1)研究協力病院は以下の基準を満たすものとする。

- ・臨床研修指定病院
- ・病院長の了解が得られている
- ・倫理委員会の承認が得られている
- ・東北地方以外の一般病院

(2)対象後期研修医は以下の基準を満たすものとする。

- ・初期研修を修了している
- ・初期研修と同じ病院で卒後3年目の後期研修を行っている(行っていた)
- ・ある程度独立して診療を行っている
- ・卒後3年目の外来新患診療録を10件^{*1}以上抽出できる

(3)外来新患診療録は以下の基準で選択する。

- ・平成22年1月から平成26年12月までの5年間に初診で内科外来を受診し、入院となった患者とする
- ・同一研修医の診療した診療録10件^{*1}以上を含む
- ・時間内と時間外の受診の別を問わない
- ・最終診断の疾患名は限定しない

^{*1}同一医師の診療録の必要件数は、目標とする標準誤差とこれまでの評価の標準誤差から算出したが、変更もありうる。

【方法】

(1)評価者 reviewers の選定
以下の基準に従い、コアメンバー委員会が選定する。

- ・内科臨床経験10年以上
- ・内科全般にわたり通じていると認められる
- ・協力病院以外の病院に勤務(対象となる後期研修医と面識がない)
- ・協力病院を訪れることが可能

(2) 研究協力病院の選定

上記基準に従い、コアメンバー委員会が選定する。

- ・ 依頼する協力病院の数は、後期研修医の数と目標件数から決める。
- ・ 評価者の拘束時間を考えて、地域性も考慮する。
- ・ 各協力病院の病院コーディネーター医師を選出し、以下の業務を依頼する。

(3) カルテピアレビューの方法

- ・ コアメンバー委員が事前に協力病院を訪れ、病院コーディネーター医師と相談の上、上記基準に従い、対象症例を選定する。
- ・ 病院コーディネーター医師と相談の上、評価者の協力病院訪問の日程を決める（評価者毎に別の日でもかまわない）
- ・ コアメンバー委員と評価者は、協力病院を訪問し、評価基準に従い評価表を作成する
- ・ 評価表はコアメンバー委員が預かり、研究事務局に保管する。

(4) 外的基準による評価表の収集方法

- ・ 病院コーディネーター医師は、上記カルテピアレビューの対象となった後期研修医のリストを作成し、連結可能匿名化する。
- ・ 病院コーディネーター医師は、上記カルテピアレビューの対象となった後期研修医の「指導医による研修医評価表（上記参照）」の記入を指導医に依頼し、回収する。
- ・ 病院コーディネーター医師は、上記カルテピアレビューの対象となった後期研修医のEPOCの結果の記入を指導医に依頼し、回収する。
- ・ 病院コーディネーター医師は、匿名化した「指導医による研修医評価表」と「EPOC結果」を研究事務局（東北大学医学教育推進センター）に郵送する。
- ・ 研究事務局（医学教育推進センター）は、

郵送されたこれらの評価票を責任もって管理する。

(5) 評価表の解析

- ・ コアメンバーは、データ解析担当者の協力のもとに、「カルテピアレビュー評価表」「指導医による研修医評価表」「EPOC 評価表」の3者の相関を調べ、基準連関妥当性を検討する。統計学的解析にはSPSSを用いる。

4. 研究成果

(1) 2013年度から2014年度にかけて、全国11病院（関東地方5病院、中部地方3病院、近畿地方1病院、中国地方2病院）を訪問して協力を依頼し、3病院（関東地方2病院、中部地方1病院）から協力の同意が得られた。

(2) 東北大学病院（2014-1-651）および全研究協力病院の倫理委員会の承認を得た。

(3) 評価者は、基盤研究（C）科学研究費（22590448）でカルテピアレビューを実施した5名（主研究の4名、予備研究の1名）に依頼した。

(4) 上述（*1）の同一医師の診療録の必要件数は、研究協力病院の研修医の担当患者数等をもとに、統計解析担当者と検討した末に、5件と設定した。

(5) 2014年4月～2015年3月に研究協力病院の内科外来を新患受診し、卒後3年目研修医13名が診療し、最終的に入院となった患者65名（1研修医あたり5名ずつ）の外来診療録を、2015年10月～2016年1月に評価者5名が病院を別々に訪れて、評価表（診療の質評価：5段階15項目）に基づいて評価した。各項目の平均値（標準偏差）を表1に示す。

表1 カルテレ뷰の項目別平均値

項目(診療の質評価)	平均値(標準偏差)
B1(医療面接・主訴)	3.06 (0.83)
B2(医療面接・その他)	2.78 (0.79)
B3(身体診察・主訴)	2.96 (0.85)
B4(身体診察・全身)	2.82 (0.87)
B5(検査オーダー)	3.58 (0.71)
B6(検査結果解釈)	3.64 (0.81)
B7(鑑別診断)	3.38 (0.90)
B8(治療)	3.65 (0.82)
B9(EBM)	3.56 (0.88)
B10(文章)	3.57 (0.75)
B11(他医への協力)	3.54 (0.83)
B12(患者への empathy)	2.98 (0.87)
B13(患者への説明)	2.44 (1.19)
B14(アウトカム)	3.67 (0.75)
B15(概略評価)	3.30 (0.78)

カルテレ뷰とは別に実施した各病院の指導医による研修医 13 名の評価(5段階 10項目)を表2に示す。

表2 指導医評価の項目別平均値

項目(診療内容の評価)	平均値(標準偏差)
P1(病歴聴取)	3.54 (0.97)
P2(身体診察)	3.46 (0.88)
P3(臨床推論)	3.62 (1.04)
P4(治療)	3.85 (0.69)
P5(医療技術)	3.85 (0.80)
P6(カルテ記載)	3.85 (1.21)
P7(患者への態度)	4.15 (0.99)
P8(他 staff との連携)	4.08 (0.95)
P9(患者アウトカム)	3.38 (0.77)
P10(概略評価)	3.69 (1.03)

カルテレ뷰評価と指導医による評価の項目別のピアソン相関係数を表3に示す。

表3A カルテレ뷰評価と指導医による評価の相関係数(1)

相関係数	P1	P2	P3	P4	P5
	病歴聴取	身体診察	臨床推論	治療	医療技術
B1(医面・主訴)	0.509	0.530	0.269	0.156	0.564
B2(医面・他)	0.205	0.243	0.129	-0.096	0.342
B3(身診・主訴)	0.121	0.132	0.065	-0.042	0.227
B4(身診・全身)	-0.150	-0.091	-0.161	-0.334	0.125
B5(検査オーダー)	-0.044	0.023	0.151	-0.244	0.154
B6(結果解釈)	0.042	-0.034	0.240	0.095	0.190
B7(鑑別診断)	0.299	0.246	0.102	0.025	0.440
B8(治療)	0.308	0.158	0.262	0.307	0.242
B9(EBM)	0.564	0.445	0.585	0.263	0.394
B10(文章)	0.189	0.252	0.331	-0.106	0.369
B11(協力)	-0.153	-0.056	0.039	-0.426	0.255
B12(empathy)	-0.262	-0.151	-0.078	-0.300	0.250
B13(説明)	-0.216	-0.202	-0.148	-0.215	-0.048
B14(アウトカム)	0.302	0.192	0.331	0.239	0.205
B15(概略評価)	0.424	0.316	0.277	0.279	0.477

表3B カルテレ뷰評価と指導医による評価の相関係数(2)

相関係数	P6	P7	P8	P9	P10
	カルテ記載	態度	連携	アウトカム	概略評価
B1(医面・主訴)	0.515	0.126	0.244	0.243	0.299
B2(医面・他)	0.322	0.050	0.200	0.171	0.079
B3(身診・主訴)	0.315	0.078	0.247	0.171	0.010
B4(身診・全身)	0.030	-0.195	-0.012	-0.101	-0.239
B5(検査オーダー)	0.179	-0.328	0.018	0.054	-0.124
B6(結果解釈)	0.281	0.124	0.346	0.300	0.160
B7(鑑別診断)	0.325	0.135	0.121	0.098	0.196
B8(治療)	0.116	0.043	0.197	0.028	0.292
B9(EBM)	0.443	0.177	0.512	0.365	0.460
B10(文章)	0.410	-0.044	0.241	0.317	0.184
B11(協力)	0.007	-0.139	0.043	0.042	-0.044
B12(empathy)	-0.124	-0.089	-0.017	-0.036	-0.101
B13(説明)	-0.306	-0.281	-0.142	-0.274	-0.223
B14(アウトカム)	0.359	0.025	0.316	0.150	0.159
B15(概略評価)	0.410	0.257	0.339	0.255	0.360

(太字は0.3以上)

研修医 13 名の「診療録評価」15 項目平均点 (5 症例平均) の平均 ± 標準偏差は 3.26 ± 0.19 (2.99-3.59)、「指導医による評価」10 項目平均点の平均 ± 標準偏差は 3.75 ± 0.81 (2.3-4.7)であった。

「診療録評価 (5 症例平均)」15 項目 (B1 ~ 15) と「指導医による評価」10 項目 (P1 ~ 10) の相関係数は、B1「医療面接が適切」vs P1「病歴聴取」($r=0.509$) B9「EBM に従っている」vs P3「臨床推論」($r=0.585$) B8「治療が適切」vs P4「治療」($r=0.307$) B15「概略評価」vs P10「概略評価」($r=0.306$) など正の相関が得られた。一方、B3「身体診察」は P2「身体診察」($r=0.132$)との相関は低く、P6「カルテ記載」($r=0.315$)との相関の方が高く、テンプレートをもとに身体診察をきちんと埋めると、身体診察をきちんととっているように「見えてしまう」可能性が示唆された。また、指導医による P7「患者・家族への態度」は、カルテレビューのどの項目ともほとんど相関しておらず、一方、B11「他医への協力を求めているか」、B12「患者への empathy は感じられるか」、B13「患者・家族への説明はなされているか」は、指導医からのどの項目とも全く相関していなかった。すなわち、このような対人関係は診療録からは読み取ることが難しい可能性が強く示唆された。

(6) 当初、EPOC 評価表との相関も計画したが、研究協力病院の 3 病院のうち 2 病院で EPOC 評価表を使用していないことが判明し、1 病院のみでは対象数が少ないことから、この解析は断念した。

(7) 以上を総括すると、我々が開発し既に信頼性が確認されている診療録ピアレビューシステムは、主要な項目 (医療面接、臨床

推論、治療、概略評価) において高い基準連関妥当性が認められた。一方、診療録では評価が困難な項目 (身体診察、患者医師関係など) も明らかとなった。この結果に基づき評価表を一部修正した上で、次に、第 3 段階の、医学教育に関わる種々の因子の診療への影響の検討を計画している。

5 . 主な発表論文等
(研究代表者、研究分担者及び連携研究者には下線)

[学会発表] (計 1 件)

亀岡淳一、大久保智哉、菊川誠、岩崎淳也、佐藤佐織、石井誠一、加賀谷豊、アウトカム評価としての診療録ピアレビューシステムの基準連関妥当性の検討 第 48 回日本医学教育学会総会、2016 年 7 月 29 ~ 30 日、大阪医科大学 (高槻市) (発表確定)

6 . 研究組織

(1) 研究代表者

亀岡 淳一 (KAMEOKA JUNICHI)
東北大学・大学院医学系研究科・准教授
研究者番号 : 30261621

(2) 研究分担者

()

研究者番号 :

(3) 連携研究者

金塚 完 (KANATSUKA HIROSHI)
東北大学・大学院医学系研究科・教授
研究者番号 : 80214435

石井 誠一 (ISHII SEIICHI)
東北大学・大学院医学系研究科・准教授
研究者番号 : 60221066

宮下 光令 (MIYASHITA MITSUNORI)
東北大学・大学院医学系研究科・教授
研究者番号 : 90301142