Geriatrics Curricula for Internal and Family Medicine Residents: Assessing Study Quality and Learning Outcomes

Huai Yong Cheng, MD, MPH
Molly Davis, BA

ABSTRACT

Background Prior reviews of geriatrics curricula for internal medicine (IM) and family medicine (FM) residents have not evaluated study quality or assessed learning objectives or specific IM or FM competencies.

Objective This review of geriatrics curricula for IM and FM residents seeks to answer 3 questions: (1) What types of learning outcomes were measured? (2) How were learning outcomes measured? and (3) What was the quality of the studies?

Methods We evaluated geriatrics curricula that reported learning objectives or competencies, teaching methods, and learning outcomes, and those that used a comparative design. We searched PubMed and 4 other data sets from 2003–2015, and assessed learning outcomes, outcome measures, and the quality of studies using the Medical Education Research Study Quality Instrument (MERSQI) and Best Evidence Medical Education (BEME) methods.

Results Fourteen studies met inclusion criteria. Most curricula were intended for IM residents in the inpatient setting; only 1 was solely dedicated to FM residents. Median duration was 1 month, and minimum geriatrics competencies covered were 4. Learning outcomes ranged from Kirkpatrick levels 1 to 3. Studies that reported effect size showed a considerable impact on attitudes and knowledge, mainly via pretests and posttests. The mean MERSQI score was 10.5 (range, 8.5–13) on a scale of 5 (lowest quality) to 18 (highest quality).

Conclusions Few geriatrics curricula for IM and FM residents that included learning outcome assessments were published recently. Overall, changes in attitudes and knowledge were sizeable, but reporting was limited to low to moderate Kirkpatrick levels. Study quality was moderate.

Introduction

The number of geriatrics fellowship–trained physicians (geriatricians), who provide appropriate care for older adults, is not expected to meet the needs of a growing aging population.[1,2] As a result, many older adults will rely on generalist physicians for their care.[1,2] Education in geriatrics for internal medicine (IM) and family medicine (FM) residents is an important step to equip future generalists and specialists to care for older adults. Equally important, geriatrics training is needed for all specialty and subspecialty residencies.[3–5]

Geriatrics education is a common topic in IM and FM literature.[6–14] Knowledge, skills, and attitudes have historically been included in IM and FM geriatrics competencies.[7,8] Recently, 26 minimum geriatrics competencies (MGCs) for IM and FM residents were developed.[15] Despite a plethora of recommendations, various educational curricula available on PubMed, and an extensive collection of geriatrics teaching materials from online medical education data sets, such as POGOe[16] and MedEdPORTAL,[17] prior reviews of geriatrics curricula[11,18] did not focus specifically on IM and FM, nor did they examine quality, so an evidence gap remains. We performed a systematic review to answer the following overarching question: What is the most effective method of teaching geriatrics to IM and FM residents? Accordingly, we conducted a review of published geriatrics curricula for IM and FM residents that asked the following questions: (1) What kinds of learning outcomes were measured? (2) How were learning outcomes measured? and (3) What was the quality of the studies?

DOI: http://dx.doi.org/10.4300/JGME-D-16-00037.1

Editor's Note: The online version of this article contains data resources and search strategy and MERSQI domain and subscale scores from 14 selected studies.

Methods

General Approach

We followed the Best Evidence Medical Education (BEME) guidelines[19] (with modifications) to conduct a systematic review of quantitative studies of geriatrics curricula for IM and FM residents. The modifications consisted of the omission of coding items 1, 7, 9, 10, 11, and 13; the omission of the formal review protocols from BEME guidelines; and the addition of the Medical Education Research Study Quality

Instrument (MERSQI), an instrument with validity evidence to assess study quality.[20–22]

Inclusion Criteria

Curriculum development for medical education consists of 6 steps: problem identification and general needs assessment, targeted needs assessment, goals and objectives, educational strategies, implementation, and evaluation and feeedback.[23] US accreditation bodies for the different levels of medical education require written curricula with fully developed educational objectives, educational methods, and evaluations.[23] To meet these requirements, we defined that information about the curriculum needed to include 3 steps: (1) learning objectives or geriatrics competencies; (2) teaching methods; and (3) evaluation of learning outcomes. The methods of evaluation (learning outcomes measurement) had to be accomplished by conducting a comparative study, which required either (1) a randomized controlled trial of the intervention versus traditional teaching (control) in 2 or more study groups; (2) a nonrandomized trial of the educational intervention versus traditional teaching (control) in 2 or more study groups; or (3) pretests and posttests or surveys before and after the intervention in a single group or 2 or more study groups. Additional inclusion criteria were that the curriculum had to (1) include geriatrics content; (2) have been designed for IM or FM residents, either alone or in combination with other disciplines; and (3) have been published in an English-language publication between 2003 and 2015.

We excluded curricula that were designed mainly for undergraduate medical education, fellows, or continuing medical education, or were designed for mixed learners (medical students, residents, or fellows); curricula primarily intended for palliative care education; or study design that was noncomparative, such as observational studies, focus groups, case series, review articles, and systematic reviews.

Study Selection and Data Collection and Extraction

The second author (M.D.) conducted a pilot study that covered geriatrics and palliative care curricula obtained via PubMed only. The study selection process is shown in the FIGURE. For potentially eligible abstracts, articles were retrieved for full review. Final decisions on which articles should be included were made by the first author (data sources and search strategies provided as online supplemental material).

The data collection sheet included the article's first author's name, journal and year of publication, study participants, study setting, rotation type, resident type, sample size, funding, learning objectives (defined as specific and measurable objectives or geriatrics competencies), teaching methods, study design and evaluation methods (learning outcomes measurements), learning outcomes, significance of learning outcomes, effect size, quality of MERSQI study items,[20–22] and strength of findings based on BEME.[19] For each study, the first author recorded (1) learning outcomes to levels 1 to 4 of Kirkpatrick criteria[19]; (2) learning objectives or competencies to the 26 MGCs and 7 domains developed by a group of national experts[15]; and (3) an effect size using Cohen's *d* (the difference of means between 2 groups or pretest and posttest in a single group divided by the standard deviation of the control group or pretest).[24] Effect sizes of 0.2, 0.5, 0.8, and 1.3 of Cohen's *d* were defined as small, medium, large, and very large, respectively.[24]

Quality Assessment

The MERSQI, developed by Cook et al[20] and Reed et al,[21,22] is used to grade the quality of medical education studies. The instrument consists of 6 domains (study design, sampling, type of data, validity of evaluation instrument, data analysis, and highest outcome). Total MERSQI scores range from 5 (lowest quality) to 18 (highest quality). Each domain is assigned a maximum score of 3. A higher score indicates higher quality of the individual or all domains combined.[20–22]

BEME methods were used to grade the strength of the findings, ranging from 1 (weakest) to 5 (strongest).[19] Grade 1 is defined as "no clear conclusions can be drawn and not significant"; Grade 2 as "results are ambiguous, but there appears to be a trend"; Grade 3 as "conclusions can probably be reached based on the results"; Grade 4 as "results are clear and very likely to be true"; and Grade 5 as "results are unequivocal."[19]

Data Synthesis and Analysis

Due to the heterogeneity of studies, we decided that it would be inappropriate to combine study results in a meta-analysis, so we instead performed a descriptive analysis. We used chi-square to compare categorical variables. Due to the small number of studies, we used nonparametric tests to compare numerical variables, with $P < .05$ considered to be statistically significant. Data entry and analysis were performed using SPSS version 22 (IBM Corp, Armonk, NY).

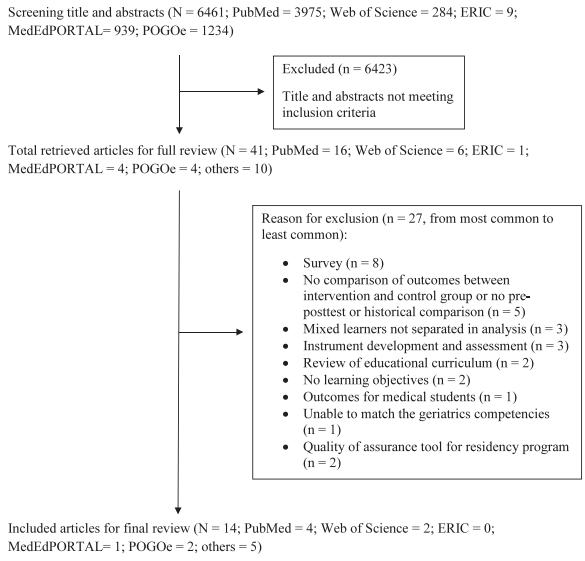

FIGURE
Flowchart of Literature Search and Selection Process (2003–2015)

Results

Study Selection and Descriptions

The study selection process is shown in the FIGURE. Fourteen studies met the inclusion criteria.[25–38]

The 14 studies are summarized in TABLE 1. Implementation, rotation sites, participants (IM or FM residents or other specialty trainees), duration, and sample size of each geriatrics curriculum varied. Seven curricula were implemented as a required geriatrics rotation.[25,27–30,32,33] Rotation sites were inpatient,[25,29,32,33,35,36] long-term care,[26,31] home visit,[29] outpatient clinic,[27,28,30,38] or some combination.[29] Three studies reported that curricula were embedded in IM rotations.[28,36,37] The duration of the geriatrics experience varied from 4 weeks[33] to 2 years.[30] All but 1 rotation (FM)[26] and 2 mixed (IM, IM-pediatrics, FM)[29,38] were designed for IM residents. Sample size varied from 28 to 180 residents,[32,37] and on average was 75. Seven studies were funded.[25,27,28,30,32,33,38]

As shown in TABLE 2, 7 of the 14 studies stated learning objectives explicitly,[28,32–37] and 7 stated them implicitly.[25–27,29–31,38] Teaching methods varied and included lecture or case-based lecture[27,29,31–34,36]; didactics[25,28,29]; web-based[30] or electronic medical record modules[26]; cues to action[38]; decision support[27,37] or simulation combined with video and case discssion[38]; grand rounds[29]; interdisciplinary team meetings or rounding[29,32]; geriatrics patient rounds[25,29,32]; nursing home rounds[31]; morning case report[28,29]; independent study projects[29]; formative feedback[35,38]; and direct observation.[25] Learning outcomes were matched to Kirkpatrick levels 1 to 4.[19]

As shown in TABLE 2, no study covered all 26 MGCs and 7 competency domains. Of the 14 studies, 3 covered only 1 of the 26 MGCs,[28,32,34] and 6 covered only 1 of the 7 competency domains.[25,28,32,34,36,38] One study covered 14 of the 26 MGCs and 6 of the 7 competency domains.[31] Median coverage of MGCs from 14 selected geriatrics curricula was 4.

Learning Outcomes

Learning outcomes are summarized in TABLE 3. All but 1 study[31] reported at least 1 statistically significant change in learning outcomes after the educational intervention. Of the 14 studies, 2 reported changes in educational experience (Kirkpatrick Level 1).[37,39] One of 2 studies showed statistically significant changes in educational experience, which were reflected in higher posttest scores.[36] Eleven studies reported changes in attitudes (Kirkpatrick Level 2a).[25–29,31,33,34,36–38] Among these, 7 studies showed statistically significant changes in learners' attitudes before and after the described curricula.[25,28,31,34,36–38] Thirteen studies reported changes in knowledge and/or skills (Kirkpatrick Level 2b).[25–33,35–38] Of these, 10 reported statistically significant changes in knowledge and/or skills.[25,29–33,35–38] Five studies reported behavioral changes in learners (Kirkpatrick Level 3),[25–28,30] and 3 of these changes were statistically significant.[25,26,30] No study reported changes in professional practice or benefits to patients (Kirkpatrick levels 4a and 4b).

We were unable to calculate effect size of the educational intervention in most studies, mainly because no standard deviations were reported. Effect size is shown in TABLE 3. In addition to effect size, sizeable changes in medical educational interventions for learning outcomes, based on the authors' comments, are shown in TABLE 3.

Methods of Analyzing Learning Outcomes

Methods of analyzing learning outcomes are summarized in TABLE 3. Designs for assessing learning outcomes also varied. All studies[25–38] used a pretest and posttest, with 29% of posttests conducted at the end of the rotation. One study used a randomized controlled trial design,[30] and another compared 2 groups without randomization.[25] The remaining studies used preintervention and postintervention in a single group.[26–29,31–38] Other assessment methods included focus groups,[26] objective structured clinical examinations,[33] encounter checklists,[30] chart reviews or audits,[25,28,30] order entry tracking,[30] direct observation,[26,28] surveys,[27] standardized patient ratings,[38]

TABLE 1
Basic Information of the Selected Geriatrics Curriculum for Internal Medicine (IM) and Family Medicine (FM) Residents

Source, y	Journal	Rotation Type	Type of Residents	Level of Residents	Sample Size
Karani et al,[33] 2004	*J Am Geriatr Soc*	It was required. 4-wk ACE rotation for all IM residents.	IM	All IM residents	35
Montagnini et al,[32] 2004	*J Palliat Med*	It was required. 1-mo geriatrics rotation at GEM unit.	IM	"Senior residents"	28
Baum and Nelson,[31] 2007	*J Am Med Dir Assoc*	1-y rotation at NH (1 half-day per month). It was not reported whether it was required rotation.	IM	PGY-1 and PGY-2	67
Westmoreland et al,[30] 2010	*J Am Geriatr Soc*	It was required. 1-mo–long ambulatory rotation.	IM and IM-peds	2 cohorts of PGY-1	96
Ahmed et al,[29] 2011	*J Am Geriatr Soc*	It was required. 1-mo geriatrics rotation at ACE unit, inpatient geriatrics and palliative care consult service, outpatient clinic and home visit.	IM and IM-peds	PGY-3 and PGY-4	32
Caton et al,[28] 2011	*J Am Geriatr Soc*	It was required. IM ambulatory clinic, month-long blocks (1 half-day continuity clinic per week).	IM, IM-peds, IM-psych	PGY-1–PGY-3 for IM; unknown for IM-peds and IM-psych	100
Litvin et al,[27] 2012	*J Am Geriatr Soc*	It was required. Ambulatory clinic, 9 mo (weekly continuity).	IM	All PGYs	100
Saffel-Shrier et al,[26] 2012	*Fam Med*	Part of FM rotation at assisted living facility, 2 y (3 half-days per month). It was not reported whether it was required.	FM	PGY-2 and PGY-3	Not reported
Wilkerson et al,[25] 2014	*J Am Geriatr Soc*	It was required. 2- to 4-wk geriatrics rotation at inpatient geriatrics consult service.	IM	Intern	42
Cumbler et al,[35] 2012	MedEdPORTAL	Duration (approximately 1 mo), inpatient geriatrics rotation. It was not reported whether it was required.	IM	PGY-1–PGY-3	Unable to calculate the number
Eskildsen et al,[36] 2012	POGOe	2-h workshop during 1-mo IM rotation. IM rotation. It was not reported whether it was required.	IM	Interns and residents	43
Brandon,[34] 2013	POGOe	1.5- to 2-h workshop, during the geriatrics rotation. It was not reported whether it was required.	IM	Residents	39

TABLE 1
Basic Information of the Selected Geriatrics Curriculum for Internal Medicine (IM) and Family Medicine (FM) Residents (continued)

Source, y	Journal	Rotation Type	Type of Residents	Level of Residents	Sample Size
Olveczky et al,[37] 2013	*JGME*	Part of 12-mo general medicine service (medicine ward) during 3-y IM residency.	IM	PGY-1–PGY-3	180
Schlaudecker et al,[38] 2013	*JGME*	3 consecutive weekly ambulatory elective, 3 half-day (afternoon) sessions, 2 to 6 residents per group, the curriculum cycle repeated monthly.	IM, FM, IM-peds, ob-gyn, PM&R	Residents	134
Summary	64% (9/14) of studies were published in geriatrics journals or geriatrics online.	50% (7/14) of studies were required geriatrics rotation. 36% (5/14) of rotation had component of inpatient setting. 29% (4/14) had component of clinic, 14% (2/14) had component of long-term care setting. 7% (1/14) rotation had component of home visit. 7% (1/14) had mixed rotation sites. The duration of geriatric experience varied from weekly to year-long (median = 1 mo).	86% (12/14) trainees were IM residents. 7% (1/14) were FM residents. 7% (1/14) had mixed residents (mainly IM, FM).	The level of residents varied from PGY-1 to PGY-4	Median = 75

Abbreviations: ACE, acute care for elderly; GEM, geriatrics evaluation and management unit; NH, nursing home; PGY, postgraduate year; IM-peds, combined internal medicine and pediatrics; IM-psych, combined internal medicine and psychiatry; MedEdPORTAL, medical education portal (https://www.mededportal.org)[16]; POGOe, Portal of Geriatrics Online Education (http://www.pogoe.org)[17]; ob-gyn, obstetrics-gynecology; PM&R, physical medicine and rehabilitation.

TABLE 2
Learning Objectives, Geriatrics Competencies and Domains, and Teaching Methods

Source, y	Statement of Learning Objectives[a]	Covered Minimum Geriatrics Competencies and Domains[b]		Teaching Methods
		Competency Items (1–26)	Competency Domains (1–7)	
Karani et al,[33] 2004	Explicit	Total = 10 (4–7, 17, 18, 21–24)	Total = 4 (2, 5–7)	Case-based learning in small group.
Montagnini et al,[32] 2004	Explicit	Total = 1 (15)	Total = 1 (4)	Bedside teaching rounds, lectures, didactic materials, multidisciplinary, and family meeting. Teaching palliative care integrated into an existing required geriatrics rotation.
Baum and Nelson,[31] 2007	Implicit	Total = 14 (1–8, 13, 14, 16–18, 23)	Total = 6 (1–5, 7)	Case-based lecture, 12 teaching sessions on different topics, nursing home rounds, teamwork with family; integrated with EOL curriculum; 12-mo longitudinal long-term care rotation.
Westmoreland et al,[30] 2010	Implicit	Total = 4 (4, 7, 23, 24)	Total = 2 (2, 7)	Web-based modules as an intervention and paper-based instruction as a control in 1-mo–long ambulatory block.
Ahmed et al,[29] 2011	Implicit	Total = 8 (4, 5, 7, 15–18, 23)	Total = 4 (2, 4, 5, 7)	Small group lectures, didactics lecture, grand rounds, patient rounds, interdisciplinary team rounding, morning case report, independent study project, 1-mo rotation.
Caton et al,[28] 2011	Explicit	Total = 1 (23)	Total = 1 (7)	Didactic presentations, system support (ie, paper and electronic cues to action), ''detailing'' (ie, review and highlight important points as outlined on the detailing sheet by an attending with a resident), an hour-long didactic session, posters with the algorithm and important messages were placed in the clinic, a monthly outpatient morning report session on the patient with fall.
Litvin et al,[27] 2012	Implicit	Total = 3 (4, 7, 23)	Total = 2 (2, 7)	Lecture, clinical decision support, and educational learning modules integrated with electronic medical records, part of the aging quality assurance project.
Saffel-Shrier et al,[26] 2012	Implicit	Total = 13 (1–4, 7–9, 13–16, 23, 24)	Total = 5 (1–4, 7)	Comprehensive geriatric assessment instrument integrated into electronic health records, interactive educational discussion on 18 competencies before each clinic session.
Wilkerson et al,[25] 2014	Implicit	Total = 3 (17–19)	Total = 1 (5)	Didactic including online reading and online learning modules, clinical rounds on inpatient geriatric consultation for 2 to 4 wk; completed ''checklist''; directly observe interns performing the Confusion Assessment Methods.
Cumbler et al,[35] 2012	Explicit	Total = 6 (1, 2, 11, 18, 19, 24)	Total = 4 (1, 3, 5, 7)	Originally used as an evaluation tool. Then, used as formative feedback prior to an inpatient rotation.

TABLE 2
Learning Objectives, Geriatrics Competencies and Domains, and Teaching Methods (continued)

Source, y	Statement of Learning Objectives[a]	Covered Minimum Geriatrics Competencies and Domains[b]		Teaching Methods
		Competency Items (1–26)	Competency Domains (1–7)	
Eskildsen et al,[36] 2012	Explicit	Total = 2 (21, 22)	Total = 1 (6)	2-h workshop during orientation and an interactive presentation once a month.
Brandon,[34] 2013	Explicit	Total = 1 (13)	Total = 1 (3)	Small group and large conference.
Olveczky et al,[37] 2013	Explicit	Total = 5 (5, 17–20)	Total = 2 (2, 5)	Bedside checklist, scripted 2-min small group educational in-service, modified several geriatric orders, drug warming system, and a decision support.
Schlaudecker et al,[38] 2013	Implicit	Total = 2 (4, 7)	Total = 1 (2)	Simulating a 10-y clinical experience with a standardized dementia patient and caregiver (daughter), lectures, videotaped longitudinal standardized patient stations, standardized patient feedback sessions, case discussion, and video review.
Summary	50% (7/14) of studies stated explicitly learning objectives	Frequency of covered MGC ≤ 5: 64% (9/14) 6–10: 21% (3/14) ≥ 11: 14% (2/14) (median = 4, varied from 1 to 14)	Frequency of covered MGC domains ≤ 2: 64% (9/14) 3–4: 21% (3/12) ≥ 5: 14% (2/14) (median = 2, varied from 1 to 6)	Teaching methods varied.

Abbreviations: EOL, end of life; MGC, minimum geriatrics competency.[15]
[a] Statement of learning objectives: "explicit" means that the author clearly listed or stated the learning objectives or described the learners will be able to—at the end of rotation or the session, "implicit" means the author implied the learning objectives in the context.
[b] The 26 minimum geriatrics competencies and 7 competency domains are obtained from Williams et al.[15]
[c] Statistically significant change in learning outcomes after the educational intervention was implemented.

management logs,[27] and a participatory process of monitoring reports.[28]

Time to assess learning outcomes is shown in TABLE 3 to demonstrate when the effect of medical intervention occurred. Response rate is also shown in TABLE 3.

Methodological Quality Assessment

Total MERSQI scores and the 6 domain scores are summarized in the online supplemental material. Mean total MERSQI score for the 14 studies was 10.5 and is shown in TABLE 2, indicating that the studies' overall methodological quality was moderate. All studies were conducted in a single institution. In 8 studies,[26,27,30,32–36] rates of follow-up with participants were either less than 50% or not reported. Validity evidence for evaluation instruments was generally poor in all studies. However, data analysis was found to be appropriate for study design and type of data in all 14 studies. The average strength of findings from the 14 studies based on BEME was 3. Eleven studies were graded as 3,[25,26,29,31–38] 1 was graded as 1, 1 as 2, and 1 as 4.[27,28,30]

Discussion

This review of IM and FM quantitative studies of geriatrics medicine curricula found that only 1 curriculum focused on FM. Learning outcomes were at low to moderate Kirkpatrick levels, yet the effects of educational intervention on attitude and knowledge were sizeable. Time of assessment via pretests and posttests (immediately or in a delayed fashion) and evidence of validity of pretests and posttests were not reported in the majority of studies. Overall study quality, as measured by MERSQI, was moderate, which was comparable to other studies of medical education research. No single study was more promising than others or stood out. The majority of educational interventions appeared feasible and could be applied or modified to apply to other institutions.

TABLE 3
Learning Outcomes and Their Measurement and Quality of Studies

Source, y	Learning Outcomes Assessment	Learning Outcomes (Kirkpatrick 1–4) and Statistical Significance[a]	Total MERSQI Scores (5–18)	Comments on Sizeable Changes of Learning Outcomes
Karani et al,[33] 2004	OSCE and pretest-posttest. It was unknown when pretest and posttest were completed. OSCE was performed at the end of 4-wk rotation. Response rate was 48%.	1, 2a, 2b,[b] (Some items in pretests-posttests were statistically significant.)	12	There were discordances between self-perceived competencies and performance (OSCE). Overall, residents were satisfied with the educational intervention. The OSCE scores were meaningful.
Montagnini et al,[32] 2004	Pretest-posttest. It was unknown when pretest and posttest were completed. Response rate was not reported.	2b[b]	9	Pretest-post changes in all 13 knowledge domains were meaningful.
Baum and Nelson,[31] 2007	Pretest-posttest. Pretest was completed in the beginning and posttest was completed at 12 mo. Response rate was 92%.	2a,[b] 2b[c] (Some items in pretests-posttests were statistically significant.)	10	Attitude (up from 3.6 to 3.7) was not meaningful. Knowledge score (up from 47% to 58%) was meaningful.
Westmoreland et al,[30] 2010	Pretest-posttest, encounter check list, chart review, order entry. It was unknown when pretest and posttest were completed. Response rate varied from 38% to 77%.	2b,[b] 3[b] (Some items in pretests-posttests were statistically significant.)	12	Effect size[b] of knowledge score is 0.5. Change of overall knowledge scores was 27.6 on a scale of 0–100, which was meaningful. Changes of chart abstract and electronic order entry scores were not meaningful.
Ahmed et al,[29] 2011	Pretest-posttest. Pretest was completed on day 1 and posttest was completed at the end of the month. Response rate was 66%.	2a, 2b[b]	10.5	Effect size of knowledge score is 0.55. Improvement of attitude was 3.9 on scale of 1–12 without statistical significance ($P = .11$), which could be meaningful.
Caton et al,[28] 2011	Pretest-posttest, observation, participation process monitoring report, and chart audit. Pretest was completed before starting the project. Posttest was completed 3 mo later. Response rate was 86%.	2a,[b] 2b, 3	11.5	Effect size of total attitude score is 0.68. There were small changes of 4 subscales of confidence. There was no comparison of observation, participation process monitoring report, and chart audit before and after the intervention.

TABLE 3
Learning Outcomes and Their Measurement and Quality of Studies (continued)

Source, y	Learning Outcomes Assessment	Learning Outcomes (Kirkpatrick 1–4) and Statistical Significance[a]	Total MERSQI Scores (5–18)	Comments on Sizeable Changes of Learning Outcomes
Litvin et al,[27] 2012	Pretest-posttest online survey, management logs. It is unknown when pretest and posttest were completed. Response rate was not reported.	2a, 2b, 3	10.5	The changes of pretest and posttest are unknown because pretest and posttest scores were not provided.
Saffel-Shrier et al,[26] 2012	Pre-posttest, direct skills observation, and focus group. It was unknown when pretest and posttest were completed. Response rate was not reported.	2a, 2b, 3[b]	11.5	Effect size of learner behavior is 2.27. Overall score change (up from 5.3 up to 7.8) was meaningful. It is unknown whether knowledge changes were meaningful because knowledge scores weren't provided
Wilkerson et al,[25] 2014	Pre-posttest and chart review, comparison between the trained and untrained groups. It was unknown when pretest and posttest were completed. Response rate was 85%.	2a,[b] 2b,[b] 3[b]	13	Changes of knowledge and self-efficacy are meaningful based on the FIGURE. Changes of 5 of 26 subscales on documenting the hazards of hospitalization were meaningful.
Cumbler et al,[35] 2012	Pretest-posttest. Pretest was completed on the orientation day. Posttest was completed following the rotation. Response rate was not reported.	2b[b]	10	Effect size of knowledge score is 1.79. Change of knowledge scores up from 70 to 91 on a scale of 0–100 was meaningful.
Eskildsen et al,[36] 2012	Pretest-posttest. Pretest was completed on the orientation day. Posttest was completed after course was over. Response rate was not reported.	1,[b] 2a,[b] 2b[b]	9	Changes of confidence up from 20 to 26 on a scale of 1–30, knowledge score change up from 4.5 to 5.7 on a scale of 8, and leaning experience up from 67% to 79% were meaningful.
Brandon,[34] 2013	Pretest-posttest. It was unknown when pretest and posttest were completed. Response rate was not reported.	2a[b]	8.5	Effect size of attitude is 1.32. Change of self-efficacy up from 3.2 to 4.6 on a scale of 1–5 was meaningful.
Olveczky et al,[37] 2013	Pretest-posttest. It was unknown when pretest and posttest were completed. Response rate varied from 60% to 91%.	2a,[b] 2b[b]	10	Changes of knowledge and self-reported ability to diagnose delirium, and awareness of delirium were meaningful.

TABLE 3
Learning Outcomes and Their Measurement and Quality of Studies (continued)

Source, y	Learning Outcomes Assessment	Learning Outcomes (Kirkpatrick 1–4) and Statistical Significance[a]	Total MERSQI Scores (5–18)	Comments on Sizeable Changes of Learning Outcomes
Schlaudecker et al,[38] 2013	Retrospective pretest-posttest, SP ratings. After each session, participants performed retrospective pretest-posttest. Response rate was 100%.	2a,[b] 2b[b]	10	Cohen's *d* is 1.19 for self-rated skills. Changes for 12 subscales of SP were meaningful.
Summary	Teaching assessments varied. Pretest-posttest were most commonly used. 64% (9/14) of studies did not report when posttests were completed. Posttests were completed at the end of rotation or immediately in 29% (4/14) and 7% (1/14) of studies, respectively. 43% (6/14) of studies did not report response rate. Mean response rate was 76%.	Frequency of Kirkpatrick level Level 1: 21% (3/14) Level 2a: 79% (11/14) Level 2b: 93% (13/14) Level 3: 36% (5/14) Level 4a: 0% Level 4b: 0% Majority of outcomes were statistically significant.	Mean total MERSQI score = 10.5	Overall, changes of attitude and knowledge from educational intervention via pretest-posttest assessments were meaningful.

Abbreviations: OSCE, objective structured clinical examination; MERSQI, Medical Education Research Study Quality Instrument[20–22]; SP, standardized patient.

[a] Levels of Kirkpatrick[19]: Level 1, participation in educational experiences; Level 2a, change of attitudes; Level 2b, change of knowledge and/or skills; Level 3, behavioral change; Level 4a, changes in professional practice; Level 4b, benefit to patients. (http://www.bemecollaboration.org).[19]

[b] Statistically significant change in learning outcomes after the educational intervention was implemented.

[c] Effect size using Cohen's *d* (ie, difference of means between 2 groups or pretest-posttest in single group divided by standard deviation of control group or pretest).[24] Effect sizes of 0.2, 0.5, 0.8, and 1.3 of Cohen's *d* were defined as small, medium, large, and very large, respectively.[24]

What Kinds of Learning Outcomes Were Measured?

We found that learning outcomes in published geriatrics curricula met only low to moderate Kirkpatrick levels.[19] However, geriatrics curricula for IM and FM residents had somewhat higher levels of Kirkpatrick criteria than other educational studies in medicine, dentistry, nursing, and pharmacy education, which have been limited to levels 1 to 2b of Kirkpatrick criteria.[39,40] Instead of pretest and posttest assessments of attitudes and knowledge, future geriatrics curriculum studies should consider assessing resident behaviors with standardized patients or patients, as well as conducting assessments more distant from the curriculum intervention. While patient outcomes represent the highest assessment level,[41–43] this is rarely feasible, due to the complexity of factors affecting patient outcomes and the large numbers of patients required. Unlike the more generous 1990s and early 2000s, little to no funding is available for this type of study. To increase the numbers and diversity of residents and to facilitate comparisons of different educational interventions, IM and FM educators should consider joining regional or national groups to study geriatrics medicine curricula.

All 14 studies except 1 reported statistically significant changes in at least 1 learning outcome.[27] Statistically significant changes in learning outcomes are not enough.[24] The effect size of a medical intervention is more important than statistical significance.[24] Unfortunately, we were able to calculate the effect size of only a few learning outcomes from several studies. The effect size from these learning outcomes was large. Nevertheless, changes in attitude and knowledge were sizeable. In other systematic reviews, the effect size or sizeable changes were not reported.[39,40] Given the small sample sizes in many of the studies in this review, these findings will need to be confirmed on a larger scale in future studies.

How Were Learning Outcomes Measured?

Learning outcomes were mainly measured by simple pretests and posttests, similar to other educational studies.[39,40] Traditionally, several methods have been used to evaluate learning outcomes.[23] Single groups with pretests and posttests are easy to implement, but because positive outcomes could be the result of natural maturation or factors unrelated to the curriculum, this method has the potential to be biased.[23] At the other end of the spectrum, controlled trials with or without randomization are complicated to implement, but are also much less likely to be biased.[23] However, nonrandomized designs are common in educational research, and are not considered to be inferior to randomized controlled trials in other areas.[44] Inherent challenges are apparent in the design and conduct of randomized controlled trial studies in medical education, such as complexity, resource-intensive evaluation,[23,44] and potential difficulty in assessing changes in professional practice and patient benefits. Also, randomization does not control for other sources of errors in education research, such as differences in implementation across settings.[44] Dividing trainees into 2 groups (ie, medical education intervention and control groups within the current structure of residency training) could pose ethical and practical challenges, and it is difficult to conduct randomized studies with residents who typically move with each rotation.[44]

What Was the Quality of Studies?

The quality of the 14 selected studies was moderate based on the MERSQI.[20–22] The MERSQI has been applied to educational research in surgery, nursing, IM residency training, complementary medicine education, Internet-based instruction, physical examination education, cultural competency, and other fields,[45] but not to geriatrics education. Its use for systematic review of geriatrics curricula, therefore, is new. Medians for total MERSQI scores and the 6 domain scores for geriatrics curricula were comparable to published medical educational studies.[45]

The strength of the 14 studies' findings was moderate (mean grade on the BEME scale was 3, on a scale of 1 to 5).[19] The BEME scale has only infrequently been reported in the literature. One systematic review of 104 articles on the effectiveness of case-based learning reported that 22 articles had grade 1; 40 had grade 2; 22 had grade 3; 13 had grade 4; and no article had grade 5.[40] The mean grade on the BEME scale was 2.2,[40] which was lower than that of the studies in this review. BEME methods[19] have not been used in geriatrics education, and are highly recommended.

In summary, this is the first systematic review to address the aforementioned 3 questions. In addition, we are the first to use MERSQI[20–22] to assess curriculum quality and BEME methods[19] to grade the strength of findings or report effect sizes in geriatrics curricula.

Our review has several limitations. First, our inclusion criteria eliminated a large number of published IM and FM geriatrics curricula, and resulted in a small number of published articles that met our criteria. Our intent was to highlight the need to measure learning outcomes in geriatrics curricula. In an era of outcome-based medical education,[41,43] evaluation of learning outcomes should be included in geriatrics curricula, and perhaps in all curricula. Second, only the first author was involved in the entirety of the project, which means that some aspects were addressed by only 1 individual. Finally, we did not use all available information resources, and we may have missed eligible studies in Scopus, CINAHL, MEDLINE, or Google Scholar.[46]

We recommend that future geriatrics curriculum studies consider assessing resident behaviors with standardized patients or patients, as well as conducting assessments more distant from the curriculum intervention. IM and FM educators should also consider expanding beyond their institution, with regional or national groups, to study geriatrics medicine curricula.

Conclusion

Between 2003 and 2015, few published geriatrics medicine curricula for FM and IM residents included assessment of learning outcomes. Overall, study quality based on MERSQI[20–22] and BEME criteria[19] was moderate, with low to moderate Kirkpatrick levels assessed. Studies often focused on only a few geriatrics competencies. Effects of educational intervention on attitude and knowledge were sizeable. The majority of geriatrics curricula reviewed can be applied or adapted to other institutions.

References

1. Institute of Medicine. *Retooling for an Aging America: Building the Health Care Workforce*. Washington, DC: National Academy of Sciences; 2008.
2. Eleazer GP, Brummel-Smith K. Commentary: aging America: meeting the needs of older Americans and the crisis in geriatrics. *Acad Med*. 2009;84(5):542–544.
3. Krain LP, Fitzgerald JT, Halter JB, et al. Geriatrics attitudes and knowledge among surgical and medical subspecialty house officers. *J Am Geriatr Soc*. 2007;55(12):2056–2060.

4. Hogan TM, Losman ED, Carpenter CR, et al. Development of geriatric competencies for emergency medicine residents using an expert consensus process. *Acad Emerg Med*. 2010;17(3):316–324.
5. Faulk CE, Lee TJ, Musick D. Implementing a multidimensional geriatric curriculum in a physical medicine and rehabilitation residency program. *Am J Phys Med Rehabil*. 2012;91(10):883–889.
6. Warshaw G, Murphy J, Buehler J, et al. Geriatric medicine training for family practice residents in the 21st century: a report from the Residency Assistance Program/Harfford Geriatrics Initiative. *Fam Med*. 2003;35(1):24–29.
7. American Geriatrics Society. Curriculum guidelines on the care of the elderly for internal medicine residency training programs. *Am J Med*. 1991;91(5):449–452.
8. Education Committee of the American Geriatrics Society. Curriculum guidelines on the care of the elderly for internal medicine residency training programs. *Am J Med*. 1997;103(4):260–262.
9. Landefeld CS, Callahan CM, Woolard N. General internal medicine and geriatrics: building a foundation to improve the training of general internists in the care of older adults. *Ann Intern Med*. 2003;139(7):609–614.
10. Hazzard WR. General internal medicine and geriatrics: collaboration to address the aging imperative can't wait. *Ann Intern Med*. 2003;139(7):597–598.
11. Thomas DC, Leipzig RM, Smith LG, et al. Improving geriatrics training in internal medicine residency programs: best practices and sustainable solutions. *Ann Intern Med*. 2003;139(7):628–634.
12. Bragg EJ, Warshaw GA. ACGME requirements for geriatrics medicine curricula in medical specialties: progress made and progress needed. *Acad Med*. 2005;80(3):279–285.
13. Chang A, Fernandez H, Cayea D, et al. Complexity in graduate medical education: a collaborative education agenda for internal medicine and geriatric medicine. *J Gen Intern Med*. 2014;29(6):940–946.
14. American Academy of Family Physicians. Recommended curriculum guideline for family medicine residents: care of older adults. http://www.aafp.org/dam/AAFP/documents/medical_education_residency/program_directors/Reprint264_Older.pdf. Accessed October 27, 2016.
15. Williams BC, Warshaw G, Fabiny AR, et al. Medicine in the 21st century: recommended essential geriatrics competencies for internal medicine and family medicine residents. *J Grad Med Educ*. 2010;2(3):373–383.
16. Association of American Medical Colleges. MededPORTAL. https://www.mededportal.org. Accessed October 27, 2016.
17. The Portal of Geriatrics Online Education (POGOe). http://www.pogoe.org. Accessed October 27, 2016.
18. Samra R, Griffiths A, Cox T, et al. Changes in medical student and doctor attitudes toward older adults after an intervention: a systematic review. *J Am Geriatr Soc*. 2013;61(7):1188–1196.
19. Association for Medical Education in Europe. Best evidence medical education. http://www.bemecollaboration.org. Accessed October 27, 2016.
20. Cook DA, Levinson AJ, Garside S. Method and reporting quality in health professions education research: a systematic review. *Med Educ*. 2011;45(3):227–238.
21. Reed DA, Cook DA, Beckman TJ, et al. Association between funding and quality of published medical education research. *JAMA*. 2007;298(9):1002–1009.
22. Reed DA, Beckman TJ, Wright SM, et al. Predictive validity evidence for medical education research study quality instrument score: quality of submissions to JGIM's Medical Education Special Issue. *J Gen Intern Med*. 2008;23(7):903–907.
23. Kern DE, Thomas PA, Hughes MT. *Curriculum Development for Medical Education: A Six-Step Approach*. 2nd ed. Baltimore, MD: The John Hopkins University Press; 2009.
24. Sullivan GM, Feinn R. Using effect size—or why the p value is not enough. *J Grad Med Educ*. 2012;4(3):279–282.
25. Wilkerson LM, Iwata I, Wilkerson MD, et al. An educational intervention to improve internal medicine interns' awareness of hazards of hospitalization in acutely ill older adults. *J Am Geriatr Soc*. 2014;62(4):727–733.
26. Saffel-Shrier S, Gunning K, Van Hala S, et al. Residency redesign to accommodate trends in geriatrics: an RC-FM variance to establish a patient-centered medical home in an assisted living facility. *Fam Med*. 2012;44(2):128–131.
27. Litvin CB, Davis KS, Moran WP, et al. The use of clinical decision-support tools to facilitate geriatric education. *J Am Geriatr Soc*. 2012;60(6):1145–1149.
28. Caton C, Wiley MK, Zhao Y, et al. Improving internal medicine residents' falls assessment and evaluation: an interdisciplinary, multistrategy program. *J Am Geriatr Soc*. 2011;59(10):1941–1946.
29. Ahmed NN, Farnie M, Dyer CB. The effect of geriatric and palliative medicine education on the knowledge and attitudes of internal medicine residents. *J Am Geriatr Soc*. 2011;59(1):143–147.
30. Westmoreland GR, Counsell SR, Tu W, et al. Web-based training in geriatrics for medical residents: a randomized controlled trial using standardized patients to assess outcomes. *J Am Geriatr Soc*. 2010;58(6):1163–1169.
31. Baum EE, Nelson KM. The effect of a 12-month longitudinal long-term care rotation on knowledge and

attitudes of internal medicine residents about geriatrics. *J Am Med Dir Assoc*. 2007;8(2):105–109.
32. Montagnini M, Varkey B, Duthie E Jr. Palliative care education integrated into a geriatrics rotation for resident physicians. *J Palliat Med*. 2004;7(5):652–659.
33. Karani R, Leipzig RM, Callahan EH, et al. An unfolding case with a linked Objective Structured Clinical Examination (OSCE): a curriculum in inpatient geriatric medicine. *J Am Geriatr Soc*. 2004;52(7):1191–1198.
34. Brandon S. Informed consent model for code status discussion. http://www.pogoe.org/productid/21684. Accessed October 27, 2016.
35. Cumbler E, Guerasio J, Youngwerth J, et al. Competency-based test of inpatient geriatric management for residents and medical students. http://www.mededportal.org/publication/9180. Accessed October 27, 2016.
36. Eskildsen M, Bonsall J, Miller A, et al. Handover and care transitions training for internal medicine residents. http://www.pogoe.org/productid/21174. Accessed October 27, 2016.
37. Olveczky D, Mattison ML, Mukamal KJ. Use of a geriatric quality initiative to educate internal medicine residents about delirium and its risk factors. *J Grad Med Educ*. 2013;5(2):309–314.
38. Schlaudecker JD, Lewis TJ, Moore I, et al. Teaching resident physicians chronic disease management: simulating a 10-year longitudinal clinical experience with a standardized dementia patient and caregiver. *J Grad Med Educ*. 2013;5(3):468–475.
39. Fatmi M, Hartling L, Hillier T, et al. The effectiveness of team-based learning on learning outcomes in health professions education: BEME Guide No. 30. *Med Teach*. 2013;35:e1608–e1624.
40. Thistlethwaite JE, Davies D, Ekeocha S, et al. The effectiveness of case-based learning in health professional education: a BEME systematic review: BEME Guide No. 23. *Med Teach*. 2012;34(6):e421–e444.
41. Chen FM, Bauchner H, Burstin H. A call for outcomes research in medical education. *Acad Med*. 2004;79(10):955–960.
42. Zafar MA, Diers T, Schauer DP, et al. Connecting resident education to patient outcomes: the evolution of a quality improvement curriculum in an internal medicine residency. *Acad Med*. 2014;89(10):1341–1347.
43. Holmboe ES, Batalden P. Achieving the desired transformation: thoughts on next steps for outcomes-based medical education. *Acad Med*. 2015;90(9):1215–1223.
44. Sullivan GM. Getting off the "gold standard": randomized controlled trials and education research. *J Grad Med Educ*. 2011;3(3):285–289.
45. Cook DA, Reed DA. Appraising the quality of medical education research methods: the Medical Education Research Study Quality Instrument and the Newcastle-Ottawa Scale-Education. *Acad Med*. 2015;90(8):1067–1076.
46. Maggio LA, Tannery NH, Kanter SL. AM last page: how to perform an effective database search. *Acad Med*. 2011;86(8):1057.

Both authors are with the University of Virginia School of Medicine. **Huai Yong Cheng, MD, MPH,** is Director of Geriatrics Clerkship, Division of General Medicine, Geriatrics, and Palliative Care Medicine and Hospital Medicine, and Associate Professor, Department of Medicine; and **Molly Davis, BA,** is a Fourth-Year Medical Student.

Funding: The authors report no external funding source for this study.

Conflict of interest: The authors declare they have no competing interests.

These study results were presented at the American Geriatrics Society annual meetings in Orlando, Florida, May 14–18, 2014; and in Washington, DC, May 14–17, 2015.

The authors would like to thank Barry Gurland, MD, Columbia University, Addeane S. Caelleigh, University of Virginia, and Editor Barbara Nordin for their assistance and helpful comments.

Corresponding author: Huai Yong Cheng, MD, MPH, University of Virginia, Division of General Medicine, Geriatrics, and Palliative Care Medicine and Hospital Medicine, PO Box 800901, Charlottesville, VA 22981, 434.924.4849, fax 434.243.9282, hyc9j@hscmail.mcc.virginia.edu

Received January 16, 2016; revisions received June 9, 2016, and July 30, 2016; accepted September 13, 2016.