補充資料文件

立法會衞生事務委員會
有關強制延續醫學教育的資料文件

背景

在2005年4月18日的衞生事務委員會會議上，議員要求政府提供補充資料，載述海外國家把延續醫學教育與醫生執業牌照的簽發掛鈎的例子、香港醫務委員會(醫委會)建議的強制延續醫學教育計劃的執行細節，以及有關延續醫學教育成效的文獻研究。本文件旨在提供所需資料。

海外國家把延續醫學教育與醫生執業牌照的簽發掛鈎的例子

美國

2. 美國有42個州份(懷俄明州要到2007年才推行強制延續醫學教育計劃)要求醫生必須符合延續醫學教育規定，方可獲得續牌。

3. 在各個州份，牌照的有效期長短不一，由一年至四年不等，而延續醫學教育周期的長短亦相應有所不同。每年規定的平均時數由12至50小時不等。

4. 未能符合延續醫學教育規定的醫生如把牌照續期，不同州份的寬限期各有不同。某些州份容許醫生補回尚欠的延續醫學教育學分，但不少州份並未設有任何寬限期。把牌照的續期與強制延續醫學教育計劃掛鈎的州份，大都透過隨機審核來監察遵行情況。

新加坡

5. 首個強制延續醫學教育周期，名為延續醫學教育資格期(CME qualifying period)，於 2003 年 1 月展開，而由 2005 年 1 月起，執業證明書的續期與符合延續醫學教育規定掛鈎。

6. 所有經正式及附有條件註冊的醫生如欲為其兩年執業證明書續期，必須在延續醫學教育資格期(對上兩年)內，至少取得 50 個延續醫學教育學分，其中 20%(即 10 個學分)須為參加與個別專科為主的延續醫學教育活動取得的基本學分，其執業證明書方可獲得續期。

7. 同樣，持有一年執業證明書的醫生，必須在延續醫學教育資格期(對上一年)內，至少取得 25 個學分，其中 20%須為基本學分，其執業證明書方可獲得續期。

8. 對於已經退休或正全時間擔任行政工作或其他職務而並非實際執業的醫生來說，他們所須符合的延續醫學教育規定較低。經新加坡醫藥理事會(Singapore Medical Council)批准，持有一年執業證明書的醫生每年必須取得 10 個學分，持有兩年執業證明書的醫生則須於兩年內取得 20 個學分。對於這些醫生，延續醫學教育不設基本學分要求，亦沒有限制在某一範疇最多可取得的學分。

9. 如未能在資格期內符合延續醫學教育規定，執業證明書在其後一年屆滿時，將不會獲得續期，而只有在執業證明書屆滿後，有關醫生才獲准開始補回尚欠的延續醫學教育學分，並申請為其執業證明書續期。因此，醫生如不遵守延續醫學教育規定，可能需要中止執業。

新西蘭

10. 醫生必須每年重新獲得認可(即更新他們的執業證明書)，方可繼續執業。不論專科醫生還是以基本醫學專業資格執業的醫

生，均須符合該項規定。有關持續專業發展[1]的規定各有不同，視乎醫生的執業範圍而定。在專責執業範圍註冊的醫生(例如專科醫生、顧問醫生、受過職業訓練的基層醫護人員)，必須符合其專業學院、協會或學會的持續專業發展規定。約有 25 個這類機構獲新西蘭醫藥理事會(Medical Council of New Zealand)認可。

11. 新西蘭醫藥理事會要求在一般執業範圍註冊的醫生，每年至少完成 50 小時的持續專業發展活動，其中每年最低限度要有一次臨牀審核，至少 20 小時教育活動，以及 10 小時同儕覆檢。

12. 醫生必須申報符合持續專業發展規定，方可重新獲得認可。每年會有 10%的醫生被檢核，以確定他們符合規定，而他們需要提供文件以資證明。

13. 倘一名醫生未獲重新認可，他會獲發臨時執業證明書，而有關個案會轉介給醫藥理事會。醫藥理事會可決定

- 在 12 個月內通過另一項檢核進行重新認可；
- 進行全面的能力覆檢；
- 施加條件限制執業範圍；
- 轉介進行健康評估；
- 建議參加補修教育計劃。

南非

14. 符合延續醫學教育的規定是與保留註冊掛鈎的。首個五年周期於 1999 年 1 月展開。執業醫生需要在五年內取得 250 個學分。南非衞生專業委員會(Health Professions Council of South Africa)或會對未能符合延續醫學教育規定的醫生施加以下規定：

- 參加補修延續醫學教育計劃；
- 參加由該委員會指定的筆試；
- 准予註冊但其執業須由該委員會作出適當監察；
- 准予註冊但其執業僅限於非臨牀範疇；

---

[1] 持續專業發展包括一系列範圍廣泛、現已被視爲專業執業一部分的技術、知識和特質，其概念較延續醫學教育更爲寬廣。

- 撤銷註冊。

15. 在聽取相關人士的意見後，委員會修訂了延續醫學教育計劃，規定醫生每 12 個月須累積 30 個延續教育單元(Continuing Education Unit)，而在累積滿 60 個延續教育單元後，必須一直保持 60 個延續教育單元。延續教育單元在有關活動舉行或結束當日起計 24 個月內有效。

16. 為方便推行有關計劃，所有醫生均會獲得 30 個學分作為起始。委員會會透過隨機抽樣審核，監察遵行延續醫學教育規定的情況。

香港醫務委員會建議的延續醫學教育計劃的細節

17. 醫委會建議採取固定周期的模式。每個延續醫學教育周期會在 1 月 1 日展開，12 月 31 日結束。

18. 醫委會建議執業證明書的有效期應由一年修訂至三年。在為期三年的延續醫學教育周期內符合延續醫學教育規定的醫生，其執業證明書可獲續期三年。此外，醫委會在三年的延續醫學教育周期結束時，可酌情向未能符合延續醫學教育規定的註冊醫生發出為期一年的執業證明書，讓他們在第四年期間，於執業時補回尚欠的延續醫學教育學分。

19. 任何醫生如在其固定的三年延續醫學教育周期中成為專科醫生，便須在該周期的餘下時間按比例計算符合有關專科醫生的延續醫學教育規定，例如：在 2006 年 1 月 1 日至 2008 年 12 月 31 日的周期內，倘一名醫生於 2007 年 7 月 1 日成為專科醫生，他必須取得專科醫生所需的 45 個延續醫學教育學分(由 2007 年 7 月 1 日至 2008 年 12 月 31 日歷時 1.5 年，所需學分因應該時期之長短按比例計算)。

20. 建議就特殊情況所作的安排如下:

- **由普通科醫生名冊的非本地名單轉移至本地名單的醫生**

他們須提供在海外參加延續醫學教育活動的證明、臨牀工作證明或其他相關證明(例如海外自願醫療服務的記錄)，而有關證明須達醫委會滿意的程度。他們的姓名一旦轉移至本地名單，便須符合有關延續醫學教育的規定。

- **在海外逗留一段時間的本地名單上的醫生**
  他們可向延續醫學教育計劃的評審者呈報所參加的延續醫學教育活動，以供評審者考慮有關活動應否獲得認可，及如獲得認可，應給予多少個延續醫學教育學分。

- **有限度註冊醫生**
  他們須符合延續醫學教育規定。專科醫生名冊上的醫生，須符合香港醫學專科學院訂明有關專科醫生的延續醫學教育規定。

- **由普通科醫生名冊除名的醫生**
  如由普通科醫生名冊除名的期間較強制延續醫學教育周期為長，或在強制延續醫學教育制度實施前除名，有關醫生須提交證據，證明在由申請把姓名重新列入普通科醫生名冊當日起計的三年期內，至少取得 30 個延續醫學教育學分。擬議的規定，適用於因未能符合延續醫學教育規定以外的原因而被醫委會從普通科醫生名冊除名的醫生。

21. 醫委會建議下列人士可獲豁免遵守延續醫學教育的規定–

- 其姓名尚未列入普通科醫生名冊第 I 部(正式註冊)的新醫科畢業生，即已獲准臨時註冊及在認可醫院或機構擔任駐院實習醫生的人士；

- 申請把其姓名列入非本地名單的人士，以及申請把保留證明書續期以便繼續列於該名單的人士；

- 基於充分理由(例如長時間患病)而未能符合延續醫學教育規定的醫生。醫委會會按個別情況考慮這些個案。不過，根據一般指引，有關醫生會獲發一年的執業證明書，

並須在重新執業的年份內至少取得 30 個延續醫學教育學分。如符合有關規定，他可獲准為其執業證明書續期，有效期至現行固定的三年延續醫學教育周期完結為止。

22. 在院校進修或擔任研究工作的有限度註冊醫生如提出申請，醫委會會按個別情況考慮豁免他們遵守有關規定。

23. 醫委會評選了延續醫學教育計劃的提供者、評審者和管理者，履行下述職能：

(a) 延續醫學教育計劃的提供者：

- 為執業醫生提供延續醫學教育活動。這些活動會由醫委會評審；
- 根據醫委會的指引對其延續醫學教育活動進行審核和頒授學分；
- 設立組織架構監察計劃的運作情況，並制訂質素保證機制；以及
- 為曾參與延續醫學教育活動的醫生提供修業證明，例如修業證明書及獲頒授的學分。

(b) 延續醫學教育計劃的評審者：

- 根據醫委會的指引，對非由醫委會認可的提供者開辦的延續醫學教育活動進行審核和頒授學分。

(c) 延續醫學教育計劃的管理者

- 為登記參加延續醫學教育計劃的執業醫生註冊；
- 核實每名參加由延續醫學教育計劃的提供者開辦的延續醫學教育活動的醫生所累積的延續醫學教育學分，並備存詳細的學分記錄；以及
- 向醫委會呈報已註冊醫生所累積的延續醫學教育學分。

延續醫學教育計劃在改善醫生執業方面的成效

24. 學習成果和執業行為的改變，往往是經過漫長時間一點一

滴產生的。與治療有所不同的是，要利用隨機試驗法[2]來評估延續醫學教育活動在維持和提升臨牀工作標準方面的成效，並不容易。儘管如此，一般認為如能參與某種形式的延續醫學教育活動，對醫護工作應有裨益。

25. 根據 Rhodes et al.所作的一項研究，延續醫學教育活動與美國外科醫學委員會重新認證考試 (American Board of Surgery Recertification Examination) 的整體表現關係密切，顯示延續醫學教育活動的參與程度偏低，本身似乎是導致考試成績不理想的風險因素。這反映延續醫學教育活動非常重要，有助醫生增進和掌握醫學知識。

26. 單是執業經驗並不足以維持專業水準。Choudhry et al.所作的一項系統檢討的結論指出，"執業時間較長的醫生可能會提供質素較低的醫護服務"。這項檢討又發現，執業年期較長的醫生和年長的醫生擁有較少基本知識；他們較容易不遵從適當的醫護標準，而且治療病人的成效也可能較為遜色。作者認為最合理的解釋是醫生的"工具箱"在他們早年受訓時製成，此後可能沒有經常更新。

27. 有關的研究載於附件[3]。

衞生福利及食物局
2005 年 10 月

[2] 隨機試驗法用以評估一項治療或一種藥物的作用。試驗的對象隨機分爲兩個組別，一組接受治療，另一組則否。觀察所得的結果差別會歸因於該項治療／藥物的作用。

[3] 研究報告只有英文本。

EDUCATION

Continuing Medical Education Activity and American Board of Surgery Examination Performance

Robert S Rhodes, MD, FACS, Thomas W Biesten, MS, Wallace P Ritchie Jr, MD, PhD, FACS, Mark A Malangoni, MD, FACS

BACKGROUND: Surgical knowledge is the basis of successful clinical problem solving, so is thought to be an important component of overall clinical ability. Continuing medical education (CME) reinforces basic knowledge and provides exposure to new knowledge within a field. Specialty board examination performance measures this knowledge but few studies have investigated a link between such performance and CME activity. This study assessed that link on the American Board of Surgery Recertification Examination.

STUDY DESIGN: The study sample comprised 278 randomly chosen applicants for the 2000 examination. Study variables included practice type, career activity, age, gender, other Board certifications, examination attempts, community size, geographic region, nationality, and ethnicity.

RESULTS: The study sample was remarkably similar to the total candidate cohort with regard to study variables. Of the 245 sample Diplomates who took the Recertification Examination, 10.2% failed. The Pass group reported 53% more total CME hours and 38% more Category I CME hours than the Fail group. The vast majority of Category I activities were surgical, clinical. Analyzed by quartiles of total CME hours, the failure rate was only 3.4% for the highest quartile but 25.8% for the lowest quartile. For Category I hours, respective failure rates were 4.8% and 19.4%. When further stratified by practice type, the failure rate of those in solo practice was 6% for those in the highest quartile of total CME hours and 37% for those in the lowest quartile. For Category I hours, the respective failure rates were 0% and 31%.

CONCLUSIONS: There is a strong relationship between CME activity and performance on the American Board of Surgery Recertification Examination. Low CME activity and practice type appear to be independent risk factors for examination failure. The relationship of these findings to patient care outcomes has important implications. (J Am Coll Surg 2003;196:604–610. © 2003 by the American College of Surgeons)

Surgical knowledge is important to successful clinical problem solving and, as such, is believed to be an important component of overall clinical surgical ability. Rapid advances in health care accentuate the need to keep current in one's specialty. The need for continued learning is emphasized by both state medical boards and specialty boards through specific requirements for continuing medical education (CME). The overall intent of such mandatory CME is to improve patient care.[1-3]

Specialty board certification examinations are well established as valid and reliable means of assessing this knowledge. Periodic recertification examinations measure both retention of basic knowledge in a discipline and the acquisition of knowledge related to subsequent advances in a specific field. CME is the time-honored mechanism for gaining and maintaining such knowledge, but relatively few studies have assessed the relationship between CME activity and examination performance. This study sought to do so in relation to the American Board of Surgery Recertification Examination in Surgery. It further sought to characterize examination performance in relation to solo practice, a characteristic previously identified by the Board as being a relatively high risk for Recertification Examination failure (American Board of Surgery, unpublished observations). The findings are then discussed in relationship to the shortcomings of CME in improving the quality of health care.

No competing interests declared.

Presented at the 88th annual American College of Surgeons Clinical Congress, October 2002, San Francisco, CA.

Received October 9, 2002; Accepted November 19, 2002.
From The American Board of Surgery, Philadelphia, PA.
Correspondence address: Robert S Rhodes, MD, FACS, 1617 JFK Blvd, Suite 860, Philadelphia, PA 19103-1847.

ISSN 1072-7515/03/$21.00
doi:10.1016/S1072-7515(03)00008-5

METHODS

Study sample

There were 1,390 approved applications for the year 2000 Recertification Examination in Surgery. Of these, there were 1,223 Diplomates who actually took the examination and their applications were abstracted for the study variables. These variables included practice type, career activity, applicant age, gender, year of initial certification, number of examination attempts, medical school, ethnicity, and other American Board of Medical Specialties' certifications. Practice type and career activity were self-reported on a checklist incorporated in the application. Practice community size was categorized by the Metropolitan Statistical Area (MSA) designation of the applicant's office address ZIP code. Geographic region was also so determined. Options for all these variables are shown in Table 1.

All applications were placed in alphabetical order and every fifth application (n = 278) was abstracted for both the number of total CME hours and Category I CME hours during the 2 years preceding application. Both types of activity were self-reported but Category I activity required accompanying documentation.

Category I CME activity was further categorized as being clinical versus nonclinical and as occurring extramural versus intramural (ie, in hospital). Clinical activity was further categorized as being surgical versus nonsurgical, and nonclinical activity was categorized as basic science versus administrative (eg, coding and billing).

Statistical analysis

Differences in distribution of demographic variables were analyzed by chi-square (for univariate measures) and by CATANOVA (for multivariate measures). Student's *t*-test was used for comparisons of age.

Distribution of CME activity was positively skewed (ie, the peak was shifted toward lesser CME activity). So the median was used to measure central tendency. Given the wide range of reported CME activity, levels of CME activity were aggregated into quartiles. The two middle quartiles encompassed relatively narrow ranges of CME activity and were further aggregated into a middle half. Differences in pass/fail rates among quartiles were analyzed by Pearson chi-square.

RESULTS

Of the 278 sample Diplomates, 245 took the Recertification Examination; 33 did not. Table 1 shows that the

Table 1. Demographics of the Study Sample (n = 245) Versus All Examinees (n = 1,223)

Characteristic	Study sample (%)	All examinees (%)
Practice type		
Solo practice (1–2 people)	38.0	36.7
Surgical group (over 2)	29.4	28.0
Multidisciplinary group (3–10)	2.9	4.0
Multidisciplinary clinic (>10)	11.0	10.8
Government (eg, military, VA)	2.9	3.7
Academic	15.5	15.8
Other	0.4	0.1
Career activity		
Clinical practice	95.9	91.2
Research	0.0	0.7
Teaching	3.3	4.4
Administration	0.4	1.0
Other (eg, emergency medicine)	0.4	1.1
Additional certification	33.9	31.6
Male gender	90.2	91.7
Ethnicity		
Asian	13.5	12.8
Black	3.7	3.7
Caucasian	78.8	78.7
Hispanic	4.1	4.2
Native American	0.0	0.7
MSA of office location (community population)		
A (> 100,000,000)	48.0	51.0
B (250,000–999,999)	22.6	22.4
C (100,000–249,999)	8.5	8.6
D (< 100,000)	0.9	2.9
Indeterminate/other	15.1	20.0
Geographic region		
Northeast	29.8	26.4
Southeast	19.2	21.3
Midwest	21.2	23.2
Southwest	13.9	10.5
West	15.1	17.3
Non-US	0.8	1.2
Medical school		
US	83.7	79.7
Foreign IMG	15.1	18.3
US IMG	1.2	2.0
Examination attempt		
First	94.7	93.0
Second	2.9	5.0
Third or more	2.4	2.0
Examination failure rate	10.2	7.7

Table 2. Median CME Activity among the 245 Sample Diplomates Who Took the Examination, by Examination Performance

	Sample examinee group (n)		
CME Activity	Total	Pass	Fail
Total	170	176	115
Category I	101	101	73

distribution of the study variables among these 245 were remarkably similar to the total year 2000 examinee cohort (n = 1,223). The p values for the distributions among these variables all exceeded 0.25 with the exception of MSA, where p exceeded 0.10. The lower p value for this variable was primarily from a higher percentage of indeterminate/other examinees among the total cohort. The mean age of the sample Diplomates was 48.0 years for both the sample and entire examinee cohort, and median ages were 47.5 and 49.0, respectively. Both groups contained essentially identical percentages of 10-year and 20-year examinees. Among the 245 sample Diplomates who took the examination, 10.2% failed. This value is slightly, but not significantly, higher than the 7.7% failure rate of the entire cohort.

Distribution of CME activity among the sample Diplomates is shown in Table 2. The Pass group reported 53% more median hours of total CME activity and 38% more median hours of Category I CME activity than the Fail group. The preponderance of Category I CME was considered clinical/surgical activity and roughly two-thirds to three-quarters of this was extramural. Only a small proportion of CME activity was administrative; no applicant reported basic science CME. Although the Fail group had fewer median hours of CME activity, the relative distribution of activity among the various subcategories did not differ significantly from that of the Pass group ($p > 0.25$).

Except as noted below, there were no significant differences between characteristics of the 245 study sample Diplomates who took the Recertification Examination and the 33 who did not. CME activity was virtually identical for both groups (eg, the median number of Category I hours for examinees was 101 compared with 102 hours for those who did not take the examination). Demographic characteristics of the two groups were also similar; the exception was that nontakers had a higher proportion of "specialists."

Table 3 shows examination performance when analyzed by quartiles of total CME hours. There is a clear relationship between CME activity and examination performance: for total CME activity, the failure rate was 3.4% for the highest CME quartile versus 25.8% for the lowest ($p < 0.000$); for Category I CME activity, the failure rate was 4.8% for the highest quartile versus 19.4% for the lowest ($p < 0.017$). The low end of the range of CME activity in the lowest quartile is below the Board's current CME requirement because some examinees had their applications approved in years before the institution of specific minimum CME requirements. There were 19 examinees who reported less than 100 total hours of CME and 15 examinees who reported less than 60 hours of Category I.

Table 3. Recertification Examination Performance Based on Quartiles of Total CME Hours and Category I CME Hours

			Pass		Fail	
CME Activity	Range of hours	Median hours	n	%	n	%
Total activity						
Highest quartile	304–5,208	452	56	96.6	2	3.4
Middle half	126–303	172	118	94.4	7	5.6
Lowest quartile	8–125	106	46	74.2	16	25.8
Category I activity						
Highest quartile	128–356	150	59	95.2	3	4.8
Middle half	73–127	101	111	91.7	10	8.3
Lowest quartile	8–72	62	50	80.6	12	19.4

Because solo practice had previously been identified as a risk factor for failure on the Recertification Examination, we also examined the interactions of practice type, CME activity, and examination performance. As noted in Table 1, solo practice comprised the largest single category of practice type, with relatively smaller percentages of Diplomates among each of the remaining nonsolo groups. Failure rates by practice type (irrespective of CME activity) were 17.2% (16 of 93) among Diplomates in solo practice, 6.9% among those in a surgical group, 0% among those in a multidisciplinary group, 7.4% among those in a multidisciplinary clinic, 0% among those in government, 5.3% among those in academia, and 0% for the one individual who listed other. The higher failure rate among those in solo practice is distinctive and consistent with previous Board findings.

Given the relatively similar low failure rates among nonsolo practice types, these groups were aggregated into a single group for purposes of CME analysis. The

aggregate failure rate of the nonsolo practice groups was 5.9% (9 of 152). This was significantly lower than the 17.2% of those in solo practice (chi-square = 6.8326, $p < 0.01$).

As a group, those in solo practice reported fewer hours of CME than those in nonsolo practice. The median hours of total and category I activity of CME among those in solo practice was 154 and 94 hours, respectively. This compared with 185 and 103 hours, respectively, among those in nonsolo practice arrangements.

Tables 4 and 5 show the interaction of CME activity, practice type, and examination performance. In general, those in solo practices had higher failure rates than those in other practice arrangements for each quartile of total and Category 1 CME activity. For total hours (Table 4), failure rate for those in solo practice and the lowest quartile was 37% (versus 16% for the lowest quartile among those in nonsolo practices). For Category I hours (Table 5), failure rates among those in the lowest quartile were 31% and 11%, respectively. Although it appears that there was not a single failure among solo practice examinees in the highest quartile of Category I activity, in fact, one of the 20 individuals was a reexaminee. Even accounting for this, the failure rate among those in solo practice in high quartiles CME activity was comparable to those in other practice arrangements.

Notably, median CME activity within each quartile was similar for those in solo practice versus those in nonsolo practice arrangements. The greatest impact of CME activity on examination performance was in the lowest quartiles of CME activity. Here, low CME activity was associated with higher failure rates in all practice types and particularly high failure rates among those in solo practice. This strongly suggests that low CME activity is an independent risk factor for poor examination performance.

The relative distribution of subcategories of CME activity (ie, extramural versus intramural) seemed similar between those in solo practice versus other practice types within either the Pass or the Fail group. But the distribution among the various subcategories yielded too few data points for further, meaningful statistical analyses.

Solo practice also correlated with type of medical school and performance on previous American Board of Surgery examinations. Again, the numbers of Diplomates in each of these variable categories were too small for further analyses.

Table 4. Comparison of Recertification Examination Performance of Those in Solo Practice Versus Nonsolo Practice, by Quartile, for Total CME Hours

Quartile	Practice type	Median hours	Pass		Fail	
			n	%	n	%
Highest	Solo	526	16	94	1	6
	Nonsolo	447	40	98	1	2
Middle half	Solo	168	42	91	4	9
	Nonsolo	177	76	96	3	4
Lowest	Solo	105	19	63	11	37
	Nonsolo	108	27	84	5	16

DISCUSSION

Surgical knowledge is important to successful problem solving and, as such, is believed to be an important component of overall clinical surgical ability. Initial certification in surgery emphasizes the acquisition of a defined standard of medical knowledge during residency training; recertification assesses the retention of this basic knowledge and familiarity with newer developments. Continuing medical education is the traditional approach to gaining and maintaining the requisite current knowledge. The presumptive link between knowledge and quality of patient care is the rationale for a required minimum CME activity by state medical boards and specialty boards for relicensure and recertification, respectively.

This study demonstrates a strong relationship between CME activity and performance on the American Board of Surgery Recertification Examination in Surgery, a valid and reliable measure of current surgical knowledge. Surgeons with relatively low levels of CME activity have a considerably greater probability of failure on that examination. These findings confirm previous Board data that identified practice type as a risk factor for examination failure and also identify low levels of CME activity as an additional, independent risk factor

Table 5. Comparison of Recertification Examination Performance of Those in Solo Practice Versus Nonsolo Practice, by Quartile, for Category I CME Hours

Quartile	Practice type	Median hours	Pass		Fail	
			n	%	n	%
Highest	Solo	153	20	100	0	0
	Nonsolo	149	39	93	3	7
Middle half	Solo	97	39	83	8	17
	Nonsolo	101	72	97	2	3
Lowest	Solo	62	18	69	8	31
	Nonsolo	63	32	89	4	11

for examination failure. The failure rates of those in the lowest quartiles of CME activity and solo practice were at least double those in the lowest quartile with other practice arrangements. Conversely, the failure rates of those in solo practice and the highest quartiles of CME activity were no different from those in other practice arrangements and comparable CME.

At least two caveats apply to these findings. First, one should not draw any conclusions about minimal or optimal levels of CME activity. Some applicants may have only reported sufficient activity to meet Board requirements (ie, 100 hours total activity, 60 hours of which must be Category I activity, during the 2 years before application). So they may not have reported all CME activity in which they participated. This caveat is particularly germane because the minimal requirement of Category I activity (60 hours) and of total activity (100 hours) both fall within the respective lowest quartile of the study findings. This is consistent with the finding that, despite the CME requirement, examination failure rates since year 2000 remain relatively high among those in solo practice.

A second caveat is to not draw any conclusions regarding the nature of the relationships among the identified risk factors. The findings may have important implications for improving both examination and practice performance, but additional research is necessary to identify other possible risk factors and to further distinguish those that may be causes from those that may be effects. Because solo practice also has a number of other correlates, practice type may only be a surrogate for some other factor(s). So the present data do not distinguish whether the higher failure rates of surgeons with little CME activity reflect an accelerated decline in examination performance over time, whether those surgeons performed less well than their peers from the outset, or whether there is some combination of these and possibly additional factors. Evidence for these possibilities comes from studies of recertification in critical care medicine by the American Board of Internal Medicine[4] and unpublished data from the American Board of Surgery.

Overall, surgeons in solo practice had less CME activity than those in other practice arrangements. It is tempting to postulate that differences in CME activity may be associated with barriers to CME participation. Yet the MSA distribution of those in solo practice was similar to that of other practice arrangements. So if barriers do exist, they may be independent of community size. For instance, those in urban solo practice might still have great difficulty arranging practice coverage to attend CME activities.

On the other hand, when considered by quartiles, those in solo practice had comparable CME activity to those in other practice arrangements. The adverse impact of low levels of CME on examination performance was largely confined to those in the lowest quartiles of CME, so other factors appear to be involved. Possibly relevant to these findings is a recent evaluation of the Maintenance of Certification Program of the Royal College of Physicians and Surgeons of Canada.[5] Here the investigators noted appreciable differences in awareness between participants and nonparticipants. This issue warrants further analysis but is beyond the immediate scope of this article.

Improving knowledge through CME is but an intermediate step toward the goal of better patient care. Consequently, the relationship of CME to improving patient care is more complex than the issues that relate CME to examination performance. Although there is evidence that performance on cognitive examinations is related to performance in practice[6] and evidence for a link between specialty board certification and improved outcomes,[7] many such studies have methodologic weaknesses.

Improving knowledge is an important first step in improving patient care, but it is increasingly recognized that participation in CME may be necessary but not sufficient to achieve this end. Indeed, the value of mandated CME activity has been questioned because of the tendency to overemphasize the quantity and underemphasize the quality of the activity. In addition, many important types of such activities often are not eligible to receive accreditation under existing requirements. Other potentially important types of activities either are not widely available or are associated with substantial barriers to participation.

These criticisms are supported by systematic reviews of the impact of continuing medical education strategies.[8] Effective change strategies included reminders, patient-mediated interventions, outreach visits, opinion leaders, and multifaceted activities. Audits with feedback and educational materials were less effective, and formal CME conferences or activities without enabling or practice reinforcing strategies had relatively little impact. The difference between effective strategies and more conventional CME strategies was an emphasis on performance change and not just on learning.[9,10]

Specific factors reported to increase the probability that CME will produce a change in practice are peer interaction,[11] commitment to change, and assessment of results.[12,13] The latter is particularly important because additional evidence suggests that improvement in care is more likely to occur with CME activity directly linking to patient care processes.[14] Such findings support the concept that CME is but one aspect of continuing physician professional development and that increasing attention needs to be focused on a broader context.[15,16] The Department of Veterans Affairs has developed an educational system for health care professionals that explicitly assesses outcomes in this broader context.[17]

The current findings may be particularly relevant to peer interaction. Specifically, the Board criterion of solo practice is one that consists of either one or two surgeons. A two- person "solo" practice might provide more opportunities than a one-person practice for cross coverage to attend extramural CME activities. On the other hand, a two-person practice still may be an insufficient "critical mass" for the feedback needed to enable a change in practice. Further distinction between one-surgeon and two-surgeon practices is needed to clarify this point.

In summary, CME appears to be important for the cognitive knowledge measured by examination performance. But increasing evidence suggests that improving patient care requires more than just a mandate for CME activity or sole emphasis on performance on a cognitive examination. Toward this end, the American Board of Medical Specialties (ABMS) has defined six general competencies inherent in medical practice. These are patient care, communication, professionalism, medical knowledge, practice-based learning and improvement, and systems-based practice. In the coming years increasing emphasis will be placed on developing measures of performance for each of these areas. For the present, surgeons should avail themselves of existing methods to evaluate knowledge deficits. One such method is the Surgical Education and Self-Assessment Program (SESAP) offered by the American College of Surgeons. Individualizing CME activity to such deficits should enhance both examination and practice performance. Achieving these goals requires further research to identify and correct the underlying factors that adversely affect examination performance and patient care.

Author contributions

Study conception and design: Rhodes, Malangoni
Acquisition of data: Rhodes
Analysis and interpretation of data: Rhodes, Biester
Drafting of manuscript: Rhodes
Critical revision: Biester, Malangoni
Statistical expertise: Rhodes, Biester
Supervision: Ritchie

REFERENCES

1. Benson JA. Certification and recertification: one approach to professional accountability. Ann Intern Med 1991;114:238–242.
2. Cruft GE, Humphreys JW, Hermann RE, Meslauskas JA. Recertification in surgery, 1980. Arch Surg 1981;116:1093–1096.
3. Norcini JJ. Recertification in medical specialties. Acad Med 1994;69:S90–S94.
4. Norcini JJ, Lipner RS. Recertification: Is there a link between take-home and proctored examinations? Acad Med 1999;74:S28–S30.
5. Tallett SE, Babitskaya G, Parboosingh J, Rothman A. The Maintenance of Certification Program: An early evaluation. Annals RCPSC 2002;35:268–273.
6. Norcini JJ, Lipner RS. The relationship between the nature of practice and performance on a cognitive examination. Acad Med 2000;75:S68–S70.
7. Sharp LK, Bashook PG, Lipsky MS, et al. Specialty Board certification and clinical outcomes: the missing link. Acad Med 2002;77:534–542.
8. Davis DA, Thomson MA, Oxman AD, Haynes RB. Changing physician performance: A systematic review of the effect of continuing medical education strategies. JAMA 1995;274:700–705.
9. Davis DA, O'Brien MJT, Freemantle N, et al. Impact of formal continuing medical education: Do conferences, workshops, rounds, and other traditional activities change physician behavior or health care outcomes? JAMA 1999;282:867–874.
10. Mazmanian PE, Davis DA. Continuing medical education and the physician as learner: Guide to the evidence. JAMA 2002; 288:1057–1060.
11. Mazmanian PE, Daffron SR, Johnson RE, et al. Information about barriers to planned change: A randomized, controlled trial involving continuing medical education lectures and commitment to change. Acad Med 1998;73:882–886.
12. Greene J. CME success stories translate into better patient care, results. AMA News: October 2, 2000.
13. Mazmanian PE, Johnson RE, Zhang A, et al. Effects of a signature on rates of change: A randomized controlled trial involving continuing education and the commitment-to-change model. Acad Med 2001;76:642–646.
14. Greene J. New approach uses CME to reduce medical errors. AMA News: February 19, 2001.
15. Wentz DK, Paulos G. Is now the time for continuing medical education to become continuing physician professional development? J Cont Educ Health Prof 2000;20:181–187.
16. Bennett NL, Davis DA, Easterling WE, et al. Continuing medical education: A new vision of the professional development of physicians. Acad Med 2000;75:1167–1172.
17. Gilman SC, Cullen RJ, Leist JC, Craft CA. Domains-based outcomes assessment of continuing medical education: The VA's model. Acad Med 2002;77:810–817.

Annex
Paper by Choudhry et al.

Annals of Internal Medicine

Systematic Review: The Relationship between Clinical Experience and Quality of Health Care

Niteesh K Choudhry, *Robert H Fletcher*, *Stephen B Soumerai*.Annals of Internal Medicine.Philadelphia: Feb 15, 2005.Vol.142,Iss.4;pg.260,14pgs

Subjects:	Physicians,Quality of service,Age,Health care delivery
Author(s):	Niteesh K Choudhry,Robert H Fletcher,Stephen B Soumerai
Document types:	Feature
Section:	*IMPROVING PATIENT CARE*
Publication title:	Annals of Internal Medicine.Philadelphia:Feb 15, 2005.Vol.142,Iss.4;pg.260,14pgs
Source type:	Periodical
ISSN/ISBN:	00034819
ProQuest document ID:	794702391
Text Word Count	6132
Document URL:	http://proquest.umi.com/pqdlink?did=794702391&sid=1&Fmt=4&clientId=53003&

Abstract (Document Summary)

Physicians with more experience are generally believed to have accumulated knowledge skills during years in practice and therefore to deliver high-quality care. Fletcher et al systematically reassess studies relating medical knowledge and health care quality to years in practice and physician age. Results show that physicians who have been in practice longer may be at risk for providing lower-quality care.

Full Text(6132 words)

[Headnote]

Background: Physicians with more experience are generally believed to have accumulated knowledge and skills during years in practice and therefore to deliver high-quality care. However, evidence suggests that there is an inverse relationship between the number of years that a physician has been in practice and the quality of care that the physician provides.

Purpose: To systematically review studies relating medical knowledge and health care quality to years in practice and physician age.
Data Sources: English-language articles in MEDLINE from 1966 to June 2004 and reference lists of retrieved articles.
Study Selection: Studies that provided empirical results about knowledge or a quality-of-care outcome and included years since graduation or physician age as explanatory variables.
Data Extraction: We categorized studies on the basis of the nature of the association between years in practice or age and performance.
Data Synthesis: Overall, 32 of the 62 (52%) evaluations reported decreasing performance with increasing years in practice for all outcomes assessed; 13 (21%) reported decreasing performance with increasing experience for some outcomes but no association for others; 2 (3%) reported that performance initially increased with increasing experience, peaked, and then decreased (concave relationship); 13 (21%) reported no association; 1 (2%) reported increasing performance with increasing years in practice for some outcomes but no association for others; and 1 (2%) reported increasing performance with increasing years in practice for all outcomes. Results did not change substantially when the analysis was restricted to studies that used the most objective outcome measures.
Limitations: Because of the lack of reliable search terms for physician experience, reports that provided relevant data may have been missed.
Conclusions: Physicians who have been in practice longer may be at risk for providing lower-quality care. Therefore, this subgroup of physicians may need quality improvement interventions.
Ann Intern Med. 2005;142:260-273. www.annals.org

Quality assurance and performance evaluation have become central issues in medicine. Care is suboptimal in many different medical conditions and clinical settings (1-6). Although delivering high-quality care is important to all clinicians, this issue may be particularly relevant to certain subgroups, such as physicians with less specialized training and those who see a smaller volume of patients (7-10).

Physicians who have been in practice for more years may also be less likely to deliver high-quality care (11, 12). Medical advances occur frequently, and the explicit knowledge that physicians possess may easily become out-of date. Therefore, although it is generally assumed that the tacit knowledge and skills accumulated by physicians during years of practice lead to superior clinical abilities (13), it is also plausible that physicians with more experience may paradoxically be less likely to provide technically appropriate care.

Few existing studies have had the specific goal of evaluating the effects of experience on the quality of medical care (11). However, length of time in clinical practice has been included as part of a set of physician characteristics that might explain variations in quality or that may be confounders of the association between quality and other factors (13-18).

The purpose of this paper is to assess the robustness of the relationship between clinical experience and quality of care by systematically reviewing empirical studies. Although we define experience as the number of years a physician has been in practice, physician age and time in practice are highly correlated (11, 19, 20); therefore, for the purposes of this paper, we

consider these variables to be interchangeable.

METHODS

We searched MEDLINE (Ovid Technologies, 1966 to June 2004; English language) for terms describing physician experience (keywords: physician age, clinician age, physician experience, clinician experience), physician demographic characteristics (keywords: physician characteristics, clinician characteristics), practice variation (subject heading: physician's practice patterns), and performance in various domains (subject headings: clinical competence, health knowledge, attitudes and practice, outcomes assessment [health care]; keywords: knowledge, guideline adherence, appropriateness, outcomes]. We retrieved potentially relevant articles and reviewed their reference lists to identify studies that our search strategy may have missed (Figure 1). We also searched our personal archives to identify additional studies. We included studies if they 1) were original reports providing empirical results; 2) measured knowledge, guideline adherence, mortality, or some other quality-of-care process or outcome; and 3) included years since graduation from medical school, years since certification, or physician age as a potential explanatory variable. We excluded studies if they described practice variation that is not known to affect quality of care (for example, assessed test-ordering behavior in clinical situations where optimal practice is unknown) or evaluated the performance of fewer than 20 physicians. For studies that examined several different end points, we included only those outcomes that are linked to knowledge or quality of care.

We used a standardized data extraction form to obtain data on study design and relevant results. We categorized studies into 4 groups on the basis of whether they evaluated knowledge (for example, knowledge of indications for blood transfusion), adherence to standards of care for diagnosis, screening, or prevention (for example, adherence to preventive care guidelines), adherence to standards of care for therapy (for example, appropriate prescribing), or health outcomes (for example, mortality). We classified the results of each study into 6 groups on the basis of the nature of the association between length of time in practice or age and performance: consistently negative, partially negative, no effect, mixed effect partially positive, and consistently positive. "Consistently negative" studies were those for which all reported outcomes demonstrated a statistically significant decrease in performance with increasing years in practice or age. "Partially negative studies showed decreasing performance with increasing experience for some outcomes and no association for others. We used similar definitions for "consistently positive" and "partially positive" studies. "Concave" studies found performance to initially improve with years in practice or age, then peak, and subsequently decrease.

We did not use formal meta-analytic techniques because the included studies used many different effect measures and some did not report parameter estimates.

Since studies based on self-reported practice may suffer from social desirability bias (21), we explored the effect of study quality on results by subcategorizing studies according to whether they measured outcomes with self-reports (that is, using surveys and interviews) or observed practice (that is, using chart audits or administrative data review). We also compared studies according to whether they performed multivariable modeling to adjust for patient and physician covariates. We use the Fisher exact test to compare the observed frequencies. We conducted all analyses with SAS, version 8.2 (SAS Institute

Inc., Gary, North Carolina).

Role of the Funding Source

The Harvard Pilgrim Health Care Foundation supported this study. It had no role in the design, conduct, or reporting of the study or in the decision to submit the manuscript for publication.

RESULTS

Fifty-nine articles that reported data on 62 groups of relevant outcomes formed the basis of our analysis. Overall, 32 of the 62 evaluations (52%) demonstrated a negative association between increasing experience and performance (that is, performance decreased as experience increased) for all outcomes assessed; 13 (21%) reported a negative association for some outcomes but no association for other outcomes; 2 (3%) reported a concave relationship (that is, performance initially increased as experience increased, then peaked, and subsequently decreased); 13 (21%) reported no association; 1 (2%) reported a positive association (that is, performance increased as experience increased) for some outcomes but no association for other outcomes; and 1 (2%) reported a positive association for all outcomes assessed (Figure 2).

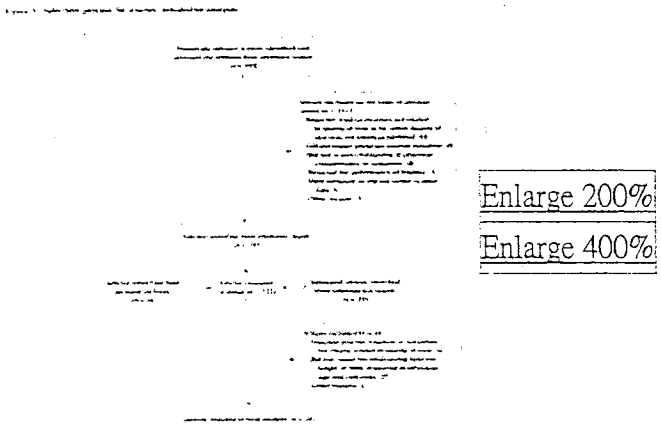
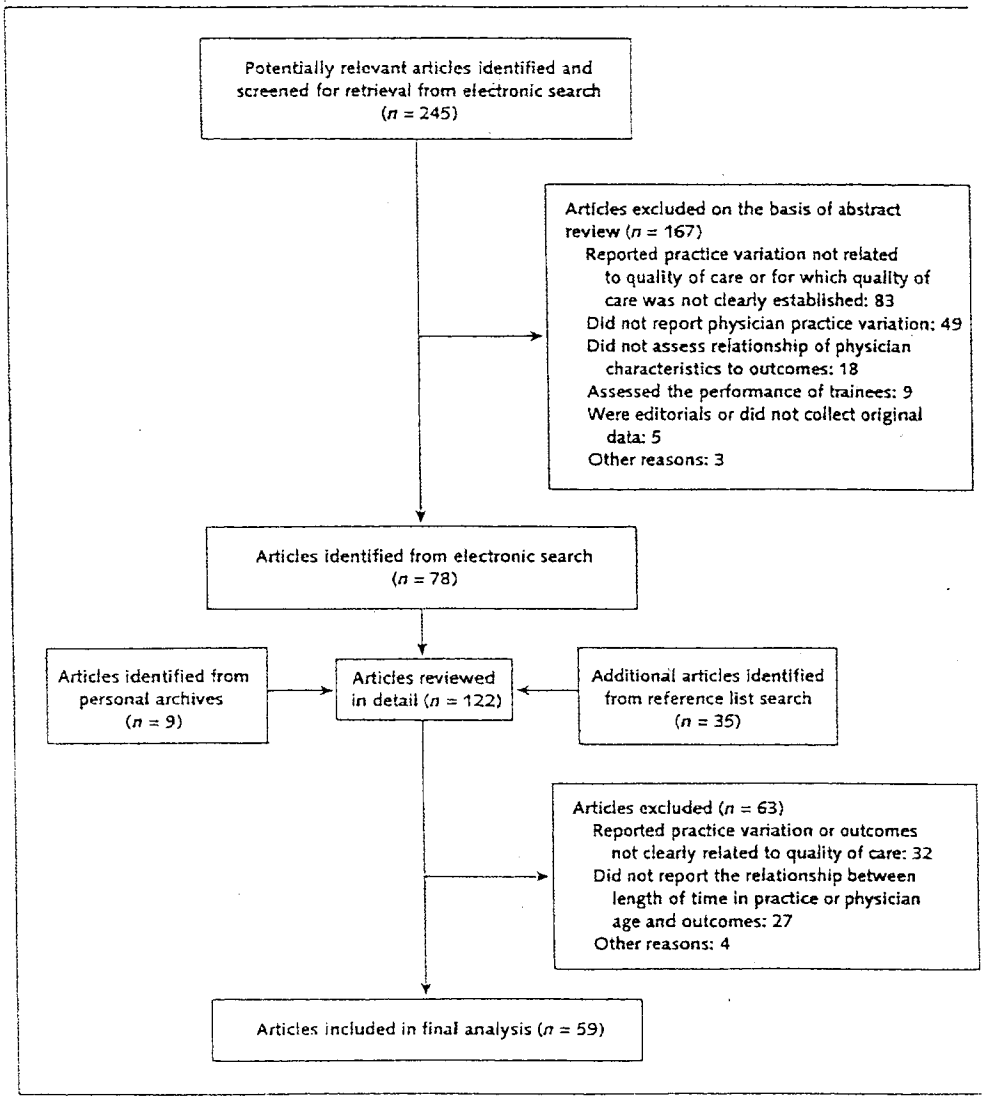


Figure 1. Selection process for studies included in analysis.

Figure 1. Selection process for studies included in analysis.



Knowledge

Twelve studies assessed the knowledge of practicing physicians, and all studies reported a negative association between knowledge and experience (Table 1). Studies by Ayanian and colleagues (7) and Salem-Schatz and colleagues (22) had large sample sizes, high response rates, and good sampling methods; used rigorous criteria to evaluate knowledge; and performed multivariate analysis.

Ayanian and colleagues (7) surveyed physicians to assess their beliefs about the survival benefit of therapies for acute myocardial infarction; the appropriate use of these therapies has been well-established in randomized, controlled trials.

Specialists were more knowledgeable than generalists; however, after adjustment for this and other variables, physicians younger than 40 years of age were more likely to correctly believe in the value of therapies that improve survival (for example, thrombolytic agents, aspirin, and β-blockers). They were also significantly less likely to believe in the value of therapies that have been disproved (for example, prophylactic lidocaine) ($P < 0.05$).

Salem-Schatz and colleagues (22) interviewed surgeons and anesthesiologists to assess their knowledge of the risks associated with and indications for the transfusion of blood products. They found a highly significant negative association between knowledge and the number of years the physicians had been in practice ($P < 0.001$).

Adherence to Standards of Practice for Diagnosis, Screening, and Prevention

Twenty-four studies have assessed the appropriateness of physician use of diagnostic and screening tests, as well as preventive health care (Table 2). Overall, 15 (63%) of these studies demonstrated that physicians in practice for more years were less likely to adhere to standards of practice in this domain.

In the largest of these studies, Czaja and colleagues (33) surveyed participants to assess their adherence to cancer screening guidelines endorsed by the American Cancer Society and the National Cancer Institute. Physicians who had graduated more than 20 years before the survey were consistently less likely to adhere to recommended practices (odds ratio, 0.62 to 0.72; $P < 0.05$).

Using more objective measures of guideline adherence, Aubin and colleagues (17) assessed the practice of 21 physicians and found that after adjustment for patient and physician covariates, younger physicians were more likely to appropriately screen for hypertension (odds ratio, 1.11 [95% CI, 1.06 to 1.15).

Several other studies provide contrary results. Streja and Rabkin (47) assessed the use of recommended preventive care measures and found that after adjustment for other physician covariates (such as specialty, practice style, and number of diabetic patients in their practice), older physicians were more likely than younger physicians to test for proteinuria (odds ratio, 2.62 [CI, 1.61 to 4.37]) and to refer their patients for screening ophthalmology assessments (odds ratio, 1.48 [CI, 1.0 to 2.18]). However, older physicians were no more likely to order a high-density lipoprotein cholesterol level test. Their analysis did not adjust for any patient variables, such as the presence of macrovascular and renal disease. Rhee (12) evaluated the performance of 454 physicians treating patients in 15 different medical and surgical diagnostic categories and found a concave relationship between years in practice and adherence to standards of practice. Physicians in practice for 6 to 15 years provided the most appropriate care, whereas physicians with more or fewer years of experience provided less appropriate care.

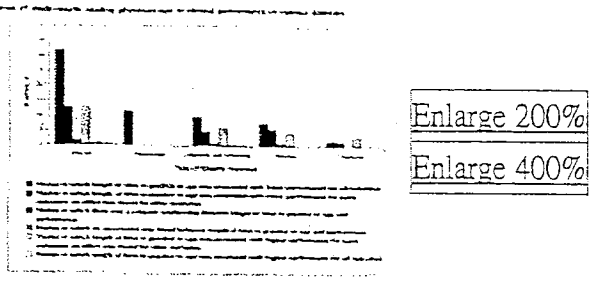
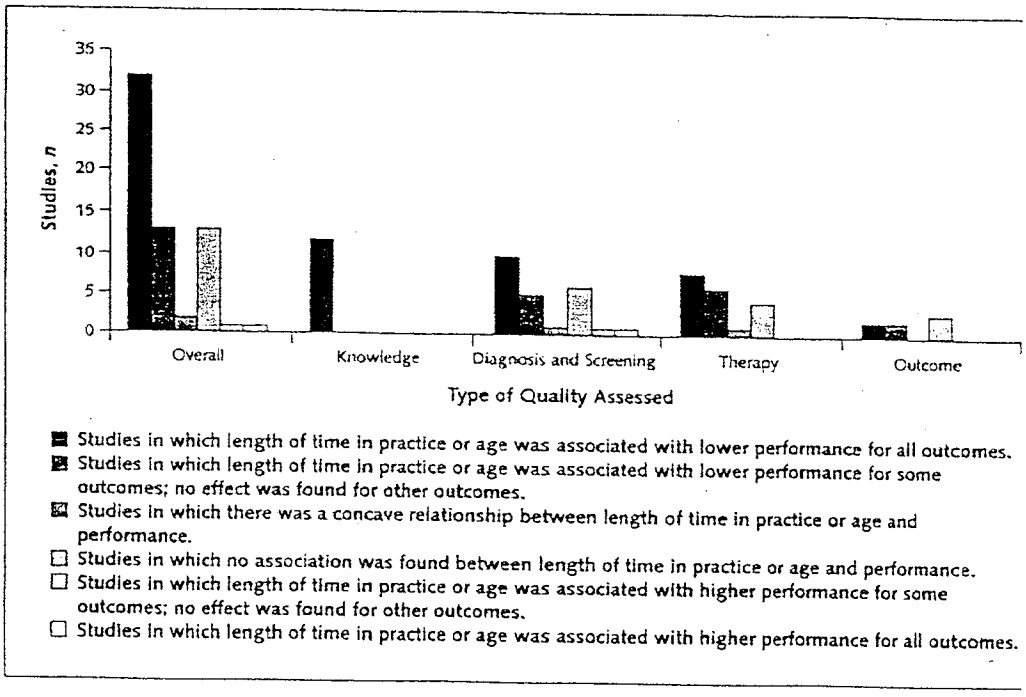


Figure 2. Distribution of study results relating physician age to clinical performance in various domains.

Figure 2. Distribution of study results relating physician age to clinical performance in various domains.



Adherence to Standards of Appropriate Therapy

Table 3 presents the 19 studies that have assessed the influence of physician age and years in practice on adherence to standards of therapy. Of these studies, 14 (74%) found a partially or consistently negative association between physician age and adherence to standards of appropriate use of therapy.

A large and well-designed study by Beaulieu and colleagues (64) examined the prescribing behavior of physicians caring for patients with stable angina. After multivariate adjustment in a hierarchical model, older physicians were significantly less likely to prescribe aspirin (odds ratio for physicians in practice for > 20 years compared with those in practice < 10 years, 0.58). Age did not affect use of β-blockers or lipid-lowering agents.

Outcomes

Seven studies present data on the relationship between number of years in practice and actual health outcomes (Table 4). The strongest of these was conducted by Norcini and colleagues (14), who analyzed mortality for 39 007 hospitalized patients with acute myocardial infarction managed by 4546 cardiologists, internists, and family practitioners. After controlling for a patient's probability of death, hospital location and practice environment, physician specialty, board certification, and the volume of patients seen, these researchers observed a 0.5% (SE, 0.27%) increase in mortality for every year since the treating physician had graduated from medical school.

Harrz and colleagues (11) specifically assessed the association between experience and mortality rates for surgeons performing cardiac artery bypass grafting. After adjustment for both patient and physician variables, they found that physicians who have been in practice longer had higher operative mortality rates ($P < 0.001$). In contrast, Burns and Wholey's (69) large study of patients hospitalized for various conditions found no difference in mortality rates for physicians of different ages, but physicians in practice for more years had longer lengths of stay even after adjustment for patients' comorbid conditions and other physician factors.

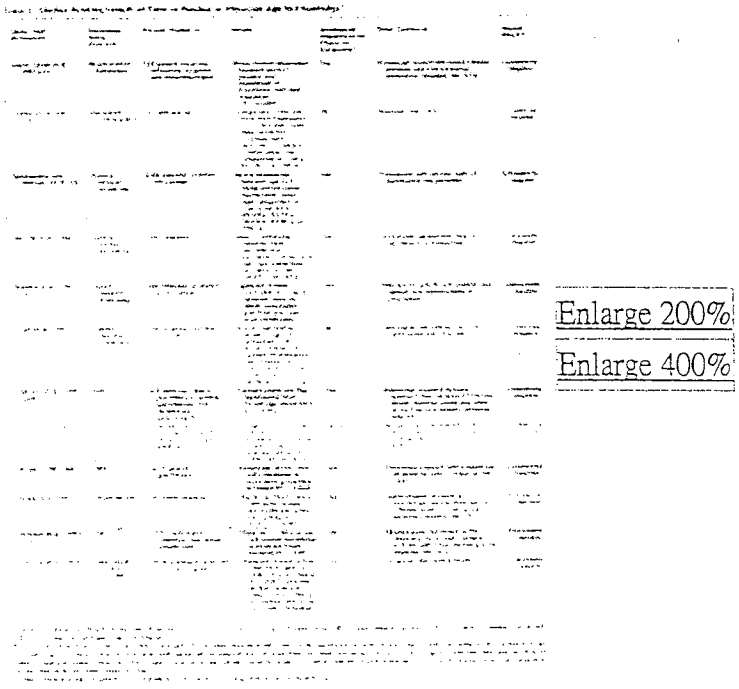


Table 1. Studies Relating Length of Time in Practice or Physician Age to Knowledge*

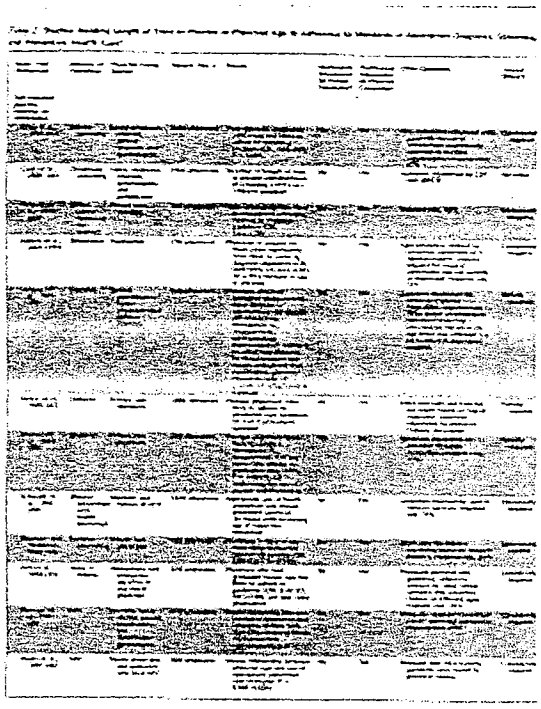
Table 1. Studies Relating Length of Time in Practice or Physician Age to Knowledge*

Study, Year (Reference)	Knowledge Being Assessed	Persons Studied, *n*	Results	Multivariate Adjustment for Physician Covariates?	Other Comments	Overall Effect?†
Salem-Schatz et al., 1990 (22)	Blood product transfusion	122 general surgeons, orthopedic surgeons, and anesthesiologists	Strong inverse association between years in practice and knowledge of transfusion risks and indications ($P = 0.0001$)	Yes	Knowledge assessment–based medical literature and NIH consensus conference; response rate, 91%	Consistently negative
Golden et al., 2001 (23)	Emergency contraception	233 pediatricians	Younger physicians and more recent graduates ($P = 0.02$) were more likely to identify FDA-approved methods of emergency contraception (age categorized as <40 y, 41–50 y, or >50 y)	No	Response rate, 24%	Consistently negative
Meskauskas and Webster, 1975 (24)	General medical knowledge	3356 internists certified ≥8 y earlier	Inverse relationship between age and ABIM recertification examination scores (age categorized as <40 y, 40–44 y, 45–49 y, 50–54 y, 54–59 y, 60–64 y, or ≥65 y)	No	Participants self-selected; tests of significance not presented	Consistently negative
Norcini et al., 1985 (25)	General medical knowledge	1947 internists	Inverse relationship between ABIM recertification examination scores and age (age categorized as <40 y, 40–49 y, 50–59 y, or ≥60 y)	No	Participants volunteered; tests of significance not presented	Consistently negative
Ramsey et al., 1991 (26)	General medical knowledge	289 internists certified 5 to 15 y earlier	Significant inverse correlation ($r = -0.3$) between score on ABIM examination questions and years since certification	Yes	Participants partially self-selected, but sample was representative of population	Consistently negative
Cruft et al., 1981 (27)	General surgical knowledge	478 surgeons certified ≥7 y earlier	Inverse relationship between age and performance on American Board of Surgery recertification examination (age categorized as 40–45 y, 46–50 y, 51–55 y, 56–60 y, or 61–73 y)	No	Participants self-selected; tests of significance not presented	Consistently negative
Gerson et al., 1991 (28)‡	HIV	473 internists, family practitioners, general practitioners, and obstetrician–gynecologists	Younger physicians had significantly more knowledge about AIDS ($P < 0.01$)	Yes	Knowledge assessed by using questions from National Center for Health Statistics survey and others devised by investigators; response rate, 63%	Consistently negative
Lewis et al., 1987 (29)‡	HIV	1000 general practitioners, family physicians, and general internists	Younger physicians had greater AIDS-related knowledge (consistent across variables assessed, although *P* values not reported)	Unclear	Measures of "competence" used were defined by group of expert clinicians at UCLA; response rate, 60%	Consistently negative
Shapiro, 1989 (30)	HIV	1271 general practitioners	Knowledge of HIV and AIDS decreased as years since graduation increased ($P = 0.008$)	Yes	Knowledge assessed with 6 questions designed by author; response rate, 70%	Consistently negative
Evans et al., 1984 (31)	Hypertension	56 family physicians	Highly significant inverse correlation between test scores and years since graduation ($r = -0.55$; $P < 0.001$)	No	Questionnaire validated to discriminate among physicians of different levels of training and specialty; response rate, 78%	Consistently negative
Ayanian et al., 1994 (7)	MI	1211 cardiologists, internists, and family practitioners	Physicians < 40 y of age had greater knowledge of evidence-based therapies ($P < 0.05$)	Yes	All physicians had served as the attending for at least 1 patient with MI within the preceding 3 mo; response rate, 61%	Consistently negative
Schroen et al., 2000 (32)	Non–small-cell lung cancer	1010 pulmonologists and thoracic surgeons	Physicians trained before 1980 more likely to underestimate survival ($P < 0.001$) and less likely to believe in value of chemotherapy in situations that have been well-established	No	Response rate, approximately 50%	Consistently negative

* ABIM = American Board of Internal Medicine; FDA = U.S. Food and Drug Administration; MI = myocardial infarction; NIH = National Institutes of Health; UCLA = University of California, Los Angeles.

† "Consistently negative" studies were those for which all outcomes demonstrated a statistically significant decrease in performance with increasing years in practice or age. "Partially negative" studies were those that showed statistically significant decreasing performance with increasing years in practice or age for some outcomes and no effect for others. Similar definitions were used for consistently positive and partially positive studies. "Concave studies" found performance to initially improve with years in practice or age then peak and subsequently decline.

‡ Also reported results on adherence to standards of diagnosis, screening, and preventive health care.



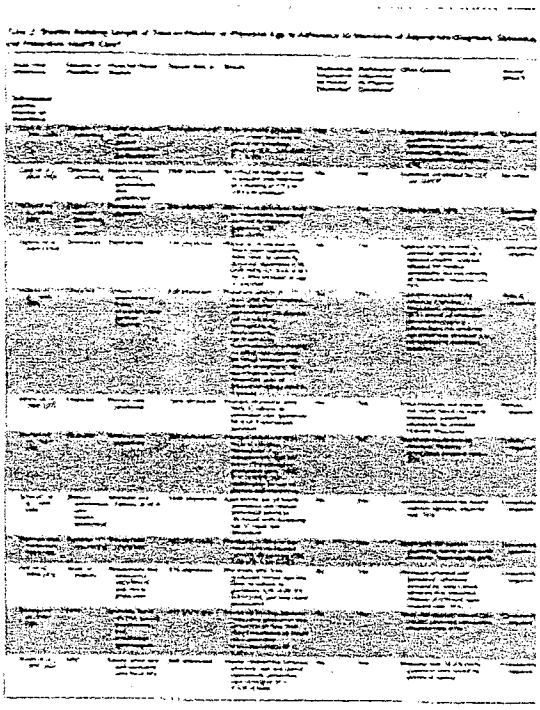
Enlarge 200%

Enlarge 400%

Table 2. Studies Relating Length of Time in Practice or Physician Age to Adherence to Standards of Appropriate Diagnosis, Screening, and Preventive Health Care*

Table 2. Studies Relating Length of Time in Practice or Physician Age to Adherence to Standards of Appropriate Diagnosis, Screening, and Preventive Health Care*

Study, Year (Reference)	Disease or Condition	Physician Group Studied	Sample Size, *n*	Results	Multivariate Adjustment for Patient Covariates?	Multivariate Adjustment for Physician Covariates?	Other Comments	Overall Effect?†
Self-reported practice (surveys or interviews)								
Czaja et al., 1994 (33)	Cancer screening	Family physicians, internists, general surgeons, and gynecologists	3436 physicians	Physicians who graduated >20 y ago less likely to adhere to screening practices (OR, 0.62–0.72; $P < 0.05$)	No	Yes	Only considered adherence with guidelines for use of interventions unanimously endorsed by multiple organizations; response rate, 67%	Consistently negative
Cook et al., 2001 (34)	Chlamydia screening	Family physicians, internists, gynecologists, and pediatricians	1600 physicians	No effect of length of time in practice and likelihood of screening (<10 y or >10 y in practice)	No	Yes	Guidelines established by CDC and USPSTF	No effect
Richards et al., 1998 (35)	Colon cancer screening for women	Primary care providers	508 physicians	Older physicians more likely to recommend screening contrary to national guidelines (OR, 3.42–3.79)	No	Yes	Response rate, 42%	Consistently negative
Epstein et al., 2001 (15)‡	Depression	Psychiatrists	278 physicians	Physicians in practice for fewer years significantly more likely to correctly diagnose depression (OR, 0.59 [95% CI, 0.43–0.81], for a 10-y increase in age or practice)	No	Yes	Appropriateness defined by consensus agreement of 4 national experts; analyses adjusted for medical comorbidity but not severity of depression; response rate, 53%	Consistently negative
Jacques et al., 1991 (36)	Diabetes	General practitioners, family physicians, and general internists	610 physicians	Physicians who had graduated more recently more likely to appropriately use glucose self-monitoring, hemoglobin A_1 measurements, ophthalmology examinations ($P < 0.001$); no effect was observed for blood pressure and weight assessment, foot examination, glycemic education; year of graduation categorized in 4 groups	No	Yes	Guidelines established by American Diabetes Association; response rate, 73%; almost all physicians performed screening interventions for which no age effect was observed (i.e., no variation in dependent variable)	Partially negative
Kenny et al., 1993 (37)	Diabetes	Primary care physicians	1434 physicians	Younger physicians more likely to adhere to preventive care guidelines for 6 of 8 procedures surveyed	No	Yes	Effect estimates not reported but results based on logistic regression; guidelines established by American Diabetes Association	Partially negative
Marrero et al., 1991 (38)	Diabetes	Primary care physicians	212 physicians	Younger physicians more likely to obtain a hemoglobin A_1 measurement (OR for every 10-y change in graduation date, 1.53; $P = 0.0017$); no relationship for use of glucose self-monitoring	No	No	Guidelines established by American Diabetes Association; response rate, 31%	Partially negative
Schwartz et al., 1991 (39)	Disease prevention and health promotion	Members and Fellows of ACP	1349 physicians	Appropriate use of health promotion and disease prevention practices decreased with increasing age (*P* value not presented)	No	Yes	Guidelines endorsed by several national agencies; response rate, 75%	Consistently negative
Sherman and Hershman, 1993 (40)	Exercise counseling	Primary care physicians	422 physicians	Physicians > 40 y of age more likely to counsel patients on exercise (OR, 3.08 [CI, 1.33–7.15])	No	Yes	Justification for exercise counseling based on research evidence; response rate, 61%	Consistently positive
Zerr et al., 1999 (41)	Fever in infants	Pediatricians and emergency department and family physicians	474 physicians	Physicians who had graduated longer ago less likely to adhere to guidelines (OR, 0.93 [CI, 0.91–0.96], per year since graduation)	No	Yes	Physicians provided with guidelines; adherence assessed by using clinical scenarios that presented children of different ages; response rate, 36%	Consistently negative
Gemson et al., 1991 (28)	HIV	Internists, family doctors, general practitioners, and obstetrician-gynecologists	473 physicians	Inverse relationship between years since graduation and adherence to New York State Department of Health AIDS prevention recommendations ($P < 0.01$)	No	Yes	Study also assessed knowledge (results presented separately); response rate, 63%	Consistently negative
Heath et al., 1997 (42)	HIV	Family physicians and specialists who treat HIV	868 physicians	Inverse relationship between physician age and use of appropriate preventive care strategies ($P < 0.001$–0.004)	No	Yes	Response rate, 38.2%–50%; guidelines were issued by provincial agency	Consistently negative



Enlarge 200%

Enlarge 400%

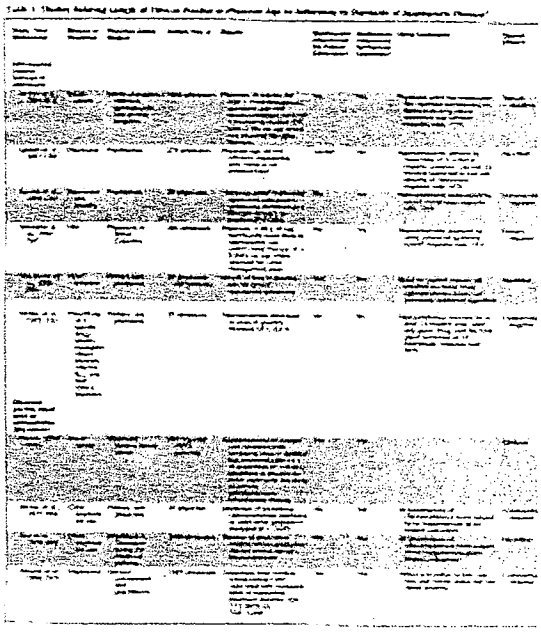
Table 2. Studies Relating Length of Time in Practice or Physician Age to Adherence to Standards of Appropriate Diagnosis, Screening, and Preventive Health Care*

Study, Year (Reference)	Disease or Condition	Physician Group Studied	Sample Size, *n*	Results	Multivariate Adjustment for Patient Covariates?	Multivariate Adjustment for Physician Covariates?	Other Comments	Overall Effect?†
Lewis et al., 1987 (29)	HIV	Primary care physicians	1000 physicians	Inverse relationship between years in practice and appropriateness of diagnostic work-up (consistent across variables assessed, although *P* values not reported)	No	Unclear	Measures of "competence" used were defined by group of expert clinicians at UCLA; study also assessed knowledge (results presented separately); response rate, 60%	Consistently negative
Skotniski et al., 1996 (43)	HIV	Primary care physicians	480 physicians	No significant association between physician age and likelihood of testing a high-risk patient	No	Unclear	Response rate, 50%; older physicians were more likely to test any patient who asked to be tested (not entirely in keeping with guidelines but unclear)	No effect
Roetzheim et al., 1991 (16)	Mammography	Primary care physicians	565 physicians	Physicians < 50 y of age were more likely than older physicians (72% vs. 49%; $P < 0.001$) to fully adhere to American Cancer Society recommendations	No	No	Response rate, 42%	Consistently negative
Ely et al., 1998 (44)	Preventive care guidelines	Primary care physicians	146 physicians	Physician age or year of graduation not associated with preventive care practices	No	Yes	Appropriateness defined by recommendations from the USPSTF; response rate, 70%	No effect
Rattay et al., 2004 (45)	Weight counseling	Pediatricians	813 physicians	Physician age not associated with frequency of weight counseling	No	Yes	Age categorized as <45 y or >45 y	No effect
Observed practice (chart audit)								
Ford et al., 1987 (46)	Breast, rectal, and small-cell lung cancer	Physicians in community hospital oncology programs	Not reported (2892 patients)	Physicians with fewer years in practice more likely to adhere to guidelines for breast and rectal cancer staging and consultation ($P < 0.01$–$P < 0.001$); no age effect was observed for small-cell lung cancer	No	No	Effect of age persisted regardless of how involved physicians were in guideline creation process	Partially negative
Streja and Rabkin, 1999 (47)	Diabetes	Primary care physicians	22 physicians (519 patients)	Physicians with >15 y experience more likely to test for proteinuria (OR, 2.62 [CI, 1.61–4.37]) and refer for ophthalmology (OR, 1.48 [CI, 1.01–2.18]) but not more likely to order an HDL cholesterol test (OR, 1.04 [CI, 0.97–1.06])	No	Yes	Did examine effect of patient characteristics on appropriate screening, but did not enter these variables into the same model as physician characteristics	Partially positive
Anis et al., 2004 (48)	Dietary and exercise counseling	Primary care physicians	38 physicians (4344 patients)	No effect of length of time in practice and likelihood of counseling	Yes	Partial	Physician covariates not significant on univariate analysis and not included in multivariate analysis	No effect
Aubin et al., 1994 (17)	Hypertension	Family physicians	21 physicians (847 patients)	Younger physicians more likely to appropriately screen for hypertension (OR, 1.11 [CI, 1.06–1.15])	Yes	Yes	Adjusted for patient age, sex, number of visits, type of visit, but not patient comorbidity; did not specify threshold for older vs. younger	Consistently negative
Hulka et al., 1976 (49)	Several (4 conditions)	Family physicians, internists, gynecologists, and pediatricians	61 physicians (1258 patients)	Physicians in practice for fewer years more likely to appropriately manage infants ($P < 0.01$). No difference observed for managing pregnancy, diabetes, or heart failure	Yes	Yes	Performance scores were developed on the basis of consensus panel discussions, all involving at least 4 family physicians as well as other physicians	Partially negative
Rhee, 1976 (12)	Several (15 diagnostic categories)	Physicians in Hawaii	454 physicians (2517 patient discharges)	Inverted "V" relationship between years in practice and adherence with standards of practice ($F = 0.01$)—physicians in practice 6–15 y provided the most appropriate care; physicians with more or fewer years in practice provided less appropriate care	No	Yes	Performance scores were developed on the basis of norms established by a "panel of physicians"; criteria not fully presented but seem to focus largely on diagnostic evaluation	Concave
Saraiya et al., 2002 (50)	Tuberculosis screening for foreign-born persons	Physicians who performed screening	491 physicians (5739 patients)	No consistent effect of number of years in practice on adherence with CDC screening recommendations	No	No	Did not have demographic data on 30% of physicians; 75% of physicians were primary care providers	No effect

* ACP = American College of Physicians; CDC = Centers for Disease Control and Prevention; HDL = high-density lipoprotein; OR = odds ratio; USPSTF = U.S. Preventive Services Task Force.

† "Consistently negative" studies were those for which all outcomes demonstrated a statistically significant decrease in performance with increasing years in practice or age. "Partially negative" studies were those that showed statistically significant decreasing performance with increasing years in practice or age for some outcomes and no effect for others. Similar definitions were used for consistently positive and partially positive studies. "Concave studies" found performance to initially improve with years in practice or age then peak and subsequently decline.

‡ Also reported results on adherence to standards of therapy—results presented separately.



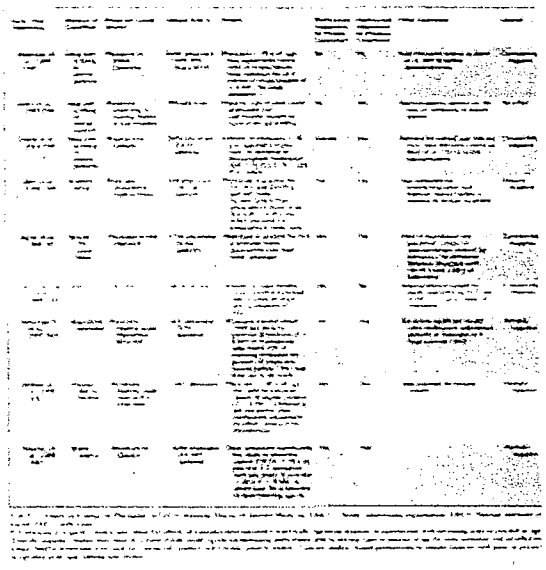
Enlarge 200%

Enlarge 400%

Table 3. Studies Relating Length of Time in Practice or Physician Age to Adherence to Standards of Appropriate Therapy*

Table 3. Studies Relating Length of Time in Practice or Physician Age to Adherence to Standards of Appropriate Therapy*

Study, Year (Reference)	Disease or Condition	Physician Group Studied	Sample Size, *n*	Results	Multivariate Adjustment for Patient Covariates?	Multivariate Adjustment for Physician Covariates?	Other Comments	Overall Effect?†
Self-reported practice (surveys or interviews)								
McFall et al., 1994 (51)	Breast cancer	Family physicians, internists, gynecologists, and general surgeons	1460 physicians	Physicians in practice for ≥20 y chose therapy less consistent with NIH recommendations for 3 of 6 treatments studied (OR, 0.56–0.78); no age effect was observed for other therapies	No	Yes	Physicians were not oncologists but reported participating in decision making about treatment and referral; response rate, 71%	Partially negative
Epstein et al., 2001 (15)	Depression	Psychiatrists	278 physicians	Physician age did not influence appropriate prescribing of an antidepressant	Limited	Yes	Appropriateness defined by consensus of 4 national experts; analyses adjusted for medical comorbidity but not severity of depression; response rate, 53%	No effect
Epstein et al., 1996 (20)	Depression and anxiety	Psychiatrists	38 physicians	"Accuracy score" (reflecting agreement with expert consensus) decreased as a function of years in practice ($P < 0.01$)	No	Yes	Appropriateness established by expert consensus; response rate, 19%	Consistently negative
Montaner et al., 1996 (52)	HIV	Physicians in British Columbia	463 physicians	Physicians < 45 y of age significantly more likely to appropriately use antiretroviral therapy ($P = 0.004$); no age effect observed for other management areas	No	Yes	Appropriateness assessed by using provincial guidelines; overall response rate, 14%	Partially negative
Roy-Byrne et al., 2002 (53)	Panic disorder	Primary care physicians	37 physicians (58 patients)	Length of time in practice did not predict appropriate prescribing	Yes	Yes	Based on patient reports (all enrolled in clinical trial); appropriateness based on previously published algorithm	No effect
Stolley et al., 1972 (13)	Prescribing of 5 specific drugs (ritalin, equagesic, chloro-mycetin, vitamin B_{12}, and oral contra-ceptives)	Primary care physicians	37 physicians	Appropriateness decreased as years in practice increased ($P < 0.01$)	No	No	Appropriateness assessed by at least 13 experts who rated any given drug, and the total panel consisted of 33 individuals; response rate, 84%	Consistently negative
Observed practice (chart audit or administrative data review)								
Hynes, 1994 (54)	Breast cancer	Physicians treating breast cancer	Not reported (3972 patients)	Appropriateness of surgical care increased with increasing years in practice but decreased after 14 y of experience ($P < 0.01$); physicians in practice for more years were less likely to provide postmastectomy rehabilitation therapy	Yes	Yes		Concave
Becker et al., 1971 (55)	Chlor-ampheni-col use	Primary care physicians	37 physicians	Likelihood of prescribing chloramphenicol increased as years since graduation increased ($P < 0.01$)	No	Yes	All prescriptions of chloramphenicol were judged to be inappropriate given limited indications	Consistently negative
Ray et al., 1976 (56)	Chlor-ampheni-col use	Physicians in Tennessee caring for Medicaid patients	3409 physicians	Number of years since medical school graduation did not predict use of chloramphenicol	No	Yes	All prescriptions of chloramphenicol were judged to be inappropriate given limited indications	No effect
Monde et al., 2002 (57)	Depression	General practitioners and psychiatrists	1527 physicians	Graduation from medical school before 1970 associated with increased odds of suboptimal treatment duration (OR, 1.12 [95% CI, 1.01–1.24])	Yes	Yes	Patient covariates include age, sex, and health status but not illness severity	Consistently negative



Enlarge 200%

Enlarge 400%

Table 3. Studies Relating Length of Time in Practice or Physician Age to Adherence to Standards of Appropriate Therapy*

Study, Year (Reference)	Disease or Condition	Physician Group Studied	Sample Size, n	Results	Multivariate Adjustment for Patient Covariates?	Multivariate Adjustment for Physician Covariates?	Other Comments	Overall
Anderson et al., 1997 (19)	Drug prescribing in elderly patients	Physicians in British Columbia	6344 physicians (819 369 drug claims)	Physicians < 45 y of age had significantly lower rates of inappropriate drug selection for all 4 classes of drugs studied (P < 0.001, for most analyses)	No	Yes	Used the same criteria as Beers et al. (58) to define appropriateness	Consistently negative
Beers et al., 1993 (58)	Drug prescribing in elderly patients	Physicians practicing in nursing homes in Los Angeles	309 physicians	Physician age or years since graduation not significantly related to appropriate prescribing	No	Yes	Appropriateness defined on the basis of consensus of expert opinion	No effect
Dhalla et al., 2002 (59)	Drug prescribing in elderly patients	Physicians in Ontario	2424 physicians (19 911 patients)	Patients of physicians > 50 y of age had a higher odds of receiving an inappropriate medication (OR, 1.14 [CI, 1.05–1.23]; P = 0.002)	Limited	Yes	Adjusted for patient age and sex only; used the same criteria as Beers et al. (58) to define appropriateness	Consistently negative
Geller et al., 1996 (60)	Hysterectomy	Physicians performing hysterectomy	339 physicians (36 104 patients)	Physicians in practice for 15–19 y and 25–29 y perform more hysterectomies than physicians in practice for 0–4 y (P < 0.05); no effect observed for physicians of other ages	Yes	Yes	Also controlled for sociodemographic and financial patient factors in addition to clinical covariates	Partially negative
Payne et al., 1984 (61)	Several (10 conditions)	Physicians in the Midwest	1135 physicians (3163 patients)	Physicians in practice for 0–9 y provided more appropriate care than other physicians	No	Yes	Tests of significance not presented; criteria for appropriateness defined by consensus; no difference between physicians with 10–19 y and >20 y of experience	Consistently negative
Sanazaro and Worth, 1985 (18)	Several	Internists	66 physicians	Number of cases treated inappropriately increased with number of years since graduation (P < 0.05)	No	No	Appropriateness judged by panel appointed by ACP and ASIM; participants were all volunteers	Consistently negative
Fehrenbach et al., 2001 (62)	Myocardial infarction	Physicians treating acute myocardial infarction	473 physicians (578 patients)	Physicians trained before 1980 less likely to prescribe β-blockers (P < 0.05); in multivariate adjustment, OR of receiving β-blocker for patients of physicians trained before 1980 was 0.66 (CI, 0.40–1.03)	Yes	Yes	Borderline-significant results after multivariate adjustment; patients all belonged to 1 large national HMO	Partially negative
Willison et al., 2000 (63)	Myocardial infarction	Physicians treating acute myocardial infarction	1452 physicians	Physicians > 50 y of age less likely to prescribe aspirin to eligible patients (P < 0.001); relationship did not persist after multivariate adjustment; no effect observed for thrombolysis	Yes	Yes	Also adjusted for hospital volume	Partially negative
Beaulieu et al., 2001 (64)	Stable angina	Physicians in Quebec	3293 physicians (11 141 patients)	Older physicians significantly less likely to prescribe aspirin (OR for < 10 y in practice, 1.7 compared with physicians in practice > 20 y; P < 0.05); no effect seen for β-blockers or lipid-lowering agents	Yes	Yes		Partially negative

* ACP = American College of Physicians; ASIM = American Society of Internal Medicine; HMO = health maintenance organization; NIH = National Institutes of Health; OR = odds ratio.

† "Consistently negative" studies were those for which all outcomes demonstrated a statistically significant decrease in performance with increasing years in practice or age. "Partially negative" studies were those that showed statistically significant decreasing performance with increasing years in practice or age for some outcomes and no effect for others. Similar definitions were used for consistently positive and partially positive studies. "Concave studies" found performance to initially improve with years in practice or age then peak and subsequently decline.

Study Quality

To determine the influence of methodologic quality on study results, we stratified the 43 reports pertaining to adherence to standards of practice on the basis of whether outcomes were assessed by using self-reported data or more objective measures (that is, use of chart audits or administrative databases). Overall, 30 (70%) of these studies demonstrated a consistently or partially negative association between length of time in practice or physician age and adherence to standards of care. While the proportion of studies that found a consistently or partially negative association was slightly larger for self-reported studies than for those studies that used objective measures (71% vs. 62%), these differences were not statistically significant ($P > 0.2$).

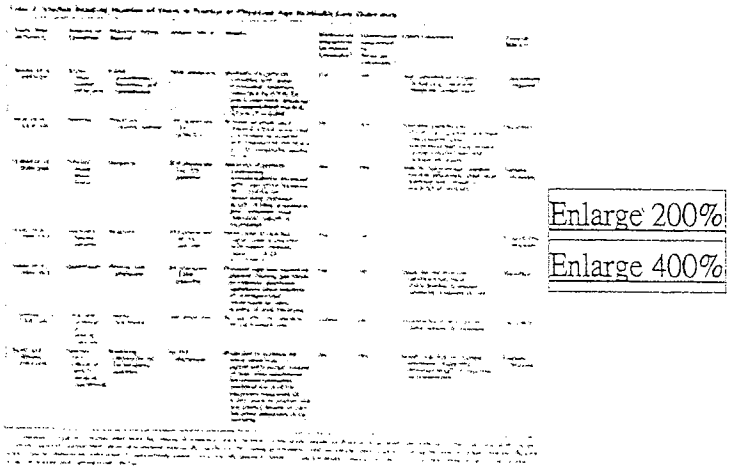


Table 4. Studies Relating Number of Years in Practice or Physician Age to Health Care Outcomes

Table 4. Studies Relating Number of Years in Practice or Physician Age to Health Care Outcomes

Study, Year (Reference)	Disease or Condition	Physician Group Studied	Sample Size, *n*	Results	Multivariate Adjustment for Patient Covariates?	Multivariate Adjustment for Physician Covariates?	Other Comments	Overall Effect?*
Norcini et al., 2000 (14)	Acute myocardial infarction	Family practitioners, internists, and cardiologists	4546 physicians	Mortality for patients admitted with acute myocardial infarction increased by 0.5% for every year since physician graduated from medical school ($P = 0.05$)	Yes	Yes	Also corrected for hospital factors (e.g., access to advanced cardiac care)	Consistently negative
Blanc et al., 2003 (65)	Asthma	Physicians treating asthma	147 physicians (317 patients)	Number of years since medical school graduation not related to patients' self-reported health status or asthma-specific quality of life	Yes	Yes	Surveyed patients and physicians separately and then linked results using hierarchical regression models; sample included very few younger physicians	No effect
O'Neill et al., 2000 (66)	Carotid endarterectomy	Surgeons	507 physicians (12 725 patients)	Mortality of patients undergoing endarterectomy increased with years since licensure ($P < 0.001$); no relationship between length of time in practice and combined "bad outcome" (death or morbidity)	Yes	Yes	Data for surgeon age available for 440 physicians; years since licensure was strongest predictor of mortality	Partially negative
Hartz et al., 1999 (11)	Coronary bypass surgery	Surgeons	275 physicians (83 547 patients)	More years in practice significantly associated with higher mortality ratios ($r = 0.22$, $P < 0.001$)	Yes	Yes		Consistently negative
Katon et al., 2000 (67)	Depression	Primary care physicians	63 physicians (1599 patients)	Physician age not related to patients having persistent or residual depressive symptoms after initiation of antidepressant medication or other quality-of-care measures	Yes	Yes	Study did not find any significant physician characteristics to explain variability in quality of care	No effect
Davidson et al., 1995 (68)	Drug prescribing in elderly patients	General practitioners	366 physicians	No age effect on mortality or hip fracture rate	Limited	No	Adjusted for patient age by using analysis of covariance	No effect
Burns and Wholey, 1991 (69)	Several (11 medical and 5 surgical conditions)	Attending physicians for hospitalized patients	54 571 discharges	Physicians in practice for more years had significantly longer lengths of stay after adjustment for patient comorbid condition for 9 of 16 diagnoses evaluated ($P < 0.05$); years in practice did not predict length of stay for other diagnoses or for mortality	Yes	Yes	Results adjusted for multiple covariates, suggesting increased length of stay may be unnecessary	Partially negative

* "Consistently negative" studies were those for which all outcomes demonstrated a statistically significant decrease in performance with increasing years in practice or age. "Partially negative" studies were those that showed statistically significant decreasing performance with increasing years in practice or age for some outcomes and no effect for others. Similar definitions were used for consistently positive and partially positive studies. "Concave studies" found performance to initially improve with years in practice or age then peak and subsequently decline.

Stratifying studies on the basis of whether they performed a multivariable analysis yielded similar results: 71% of the studies that adjusted for patient covariates found a consistently or partially negative association compared with 74% of studies that did not adjust for these factors, and 68% of the studies that adjusted for physician covariates found a consistently or partially negative association compared with 67% of studies that did not.

DISCUSSION

Although based on heterogeneous studies, our systematic review of empirical studies evaluating the relationship between clinical experience and performance suggests that physicians who have been in practice for more years and older physician possess less factual knowledge, are less likely to adhere to appropriate standards of care, and may also have poorer patient outcomes. These effects seem to persist in those studies that adjusted for other known predictors of quality, such as patient comorbidity and physician volume or specialization. The results are somewhat paradoxical since it is generally assumed that clinical experience enhances knowledge and skill and, therefore, leads to better patient care.

Our findings have many possible explanations. Perhaps most plausible is that physicians' "toolkits" are created during training and may not be updated regularly (70). Older physicians seem less likely to adopt newly proven therapies (71, 72) and may be less receptive to new standards of care (73). In addition, practice innovations that involve theoretical shifts, such as the use of less aggressive surgical therapy for early-stage breast cancer or protocols for reducing length of stay, may be harder to incorporate into the practice of physicians who have trained long ago than innovations that add a procedure or technique consistent with a physician's preexisting knowledge (74).

Our findings may also reflect the substantial environmental changes that have occurred in medicine over the past several decades; evidence-based medicine has been widely adopted, and quality assurance techniques, such as disease management and performance evaluation, are frequently used. More experienced physicians may have less familiarity with these strategies and may be less accepting of them. Given this, our results may represent a cohort effect; that is, when the current generation of more recently trained physicians has been in practice for a longer time, there may be smaller differences between their practice and those of their younger colleagues than our data would suggest.

Our study has several limitations. First, although we attempted to systematically review the literature on the association between number of years in practice or physician age and performance, our search strategy may have missed reports. This reflects the limited attention to this issue and the lack of consistent search terms to identify clinical experience. In addition, studies that were specifically designed to assess the relationship between experience and performance but found no association may have been less likely to be submitted or accepted for publication, and published studies that included number of years in practice or age among other physician characteristics may not have presented non-statistically significant results for these particular variables. Therefore, while we have no reason to suspect that we were more likely to identify studies showing decreasing performance with age, our findings are still potentially subject to an under-reporting bias.

Second, few reports included in this review were designed to specifically evaluate length of time in practice as their primary characteristic of interest. Consequently, our results may have been due to chance arising from multiple testing. However, we believe this is unlikely given the relative consistency of the results in several different domains, their "dose-response" relationship, and their overall plausibility. Moreover, restricting our analysis to the 32 studies that considered a broader set of physician characteristics, including number of years in practice or age as the focus of their investigation (that is, excluding those studies that considered physician age or number of years in practice only as confounders), does not change our results: 21 of the 32 (66%) studies reported a consistently or partially negative association between physician age and performance, whereas only 1 study demonstrated a partially positive association.

Third, disagreements may exist between clinical practice guidelines (33), and, thus, establishing appropriate norms may be difficult. As a result, assessing performance on the basis of guideline adherence may not reliably assess health care quality. Despite this, some studies included in our review used norms that had been adopted by several professional associations and that consequently reflect widely accepted standards of practice. Even for these studies, we observed age effects.

Finally, length of time in practice may be associated with other dimensions of quality that are not captured by the outcome measures that we evaluated. While we identified studies that assessed various conditions and aspects of performance, the relationship between age and performance may be different for other diseases and outcomes. For example, older physicians may be more effective at delivering the humanistic, rather than the technical, aspects of medical care. If this were true, one would expect that the patients of older physicians would report higher satisfaction, which has been demonstrated in some studies (75, 76) but not others (77, 78). Alternatively, physicians who have been in practice for a longer time may have better clinical judgment and may thus provide better care in complex cases or may be better diagnosticians. These outcomes have not been rigorously assessed.

Despite these limitations, our results are troubling. Although it is difficult to draw firm conclusions about the performance of older physicians in managing specific conditions or clinical scenarios, our results do suggest that older physicians may need quality improvement interventions that are generally applicable to all physicians. In addition, the requirements that are imposed on physicians to keep up to date and to demonstrate continuing competence should be further considered. Widely adopted continuing medical education techniques, such as the distribution of printed materials and lectures, are largely ineffective even in experimental conditions (79). Our results reinforce this. Moreover, many experienced physicians are exempt from the recertification requirements to which their more recently trained colleagues must adhere. For example, the American Board of Internal Medicine only requires physicians who received initial Board certification in or after 1990 to appear for periodic recertification examinations.

Our results also have implications for further research. The link between experience and performance should be further evaluated with studies that are designed a priori to specifically measure this association. These studies should use objective and widely accepted measures of performance; should be disease- or process-specific; and should be replicated for physicians of different specialties, demographic characteristics (such as sex), and different environment practices. The effect of age for physicians who routinely collaborate with other physicians, who frequently engage in evidence-based discussions, or whose practices are influenced by disease management, performance feedback, and computerized reminder systems may be different from that for physicians who practice in relative isolation or in more traditional settings.

An optimal study would follow a particular cohort of physicians over time. However, this is not practical and may be confounded by other secular trends in health care provision. Alternative designs would be similar to those of the highest quality included in our review and would adequately control for patient comorbidity, other physician factors, and the clustering of patients within physicians. These studies should also model the nature of the relationship between experience and performance since performance may improve during the initial phases of independent practice, plateau for some period of time, and then decrease. Finally, the ability of behavior change strategies to reduce the disparities in quality created by physician age should be evaluated in well-controlled clinical trials.

In summary, our results suggest that physicians with more experience may paradoxically be at risk for providing lower-quality care. The extent, magnitude, and nature of these results must be clarified, and added attention should be given to this subgroup of physicians who may need quality improvement interventions.

[Sidebar]

Improving Patient Care is a special section within Annals supported in part by the U.S. Department of Health and Human Services (HHS) Agency for Healthcare Research and Quality (AHRQ). The opinions expressed in this article are those of the authors and do not represent the position or endorsement of AHRQ or HHS.

[Sidebar]

Context

While "practice makes perfect" in some situations, physicians' knowledge and performance may decline with the passage of time.

Contribution

Of 62 published studies that measured physician knowledge or quality of care and described time since medical school graduation or age, more than half suggested that physician performance declined over time for all outcomes measured. Only 1 study showed improved performance for all outcomes measured.

Implications

This review should provoke careful study of the relationship of physician experience and the quality of care. It also raises concerns about the adequacy of continuing professional education in medicine.

-The Editors

[Reference]

Reference

1. McGlynn EA, Asch SM, Adams J, Keesey J, Hicks J, DeCristofaro A, et al. The quality of health care delivered to adults in the United States. N Engl J Med. 2003:348:2635-45. [PMID: 12826639]

2. Schuster MA, McGlynn EA, Brook RH. How good is the quality of health care in the United States? Milbank Q. 1998;76:517-63, 509. [PMID: 9879302]

3. Jha AK, Perlin JB, Kizer KW, Dudley RA. Effect of the transformation of the Veterans Affairs Health Care System on the quality of care. N Engl J Med. 2003:348:2218-27. [PMID: 12773650]

4. Jencks SF, Guerdon T, Burwen DR, Fleming B, Houck PM, Kussmaul AE, et al. Quality of medical care delivered to Medicare beneficiaries: A profile at state and national levels. JAMA.. 2000;284:1670-6. [PMID: 11015797]

5. Thompson JW, Ryan KW, Pinidiya SD, Bost JE. Quality of care for children in commercial and Medicaid managed care. JAMA. 2003:290:1486-93. [PMID: 13129989]

6. Chassin MR, Galvin RW. The urgent need to improve health care quality. Institute of Medicine National Roundtable on Health Care Quality. JAMA. 1998:280:1000-5. [PMID: 9749483]

7. Ayanian JZ, Hauptman PJ, Guadagnoli E, Antman EM, Pashos CL, McNeil BJ. Knowledge and practices of generalist and specialist physicians regarding drug therapy for acute myocardial infarction. N Engl J Med. 1994;331:1136-42. [PMID 7935639]
8. Halm EA, Lee C, Chassin MR. Is volume related to outcome in Health care? A systematic review and methodologic critique of the literature. Ann Intern Med. 2002;137:511-20. [PMID: 12230353]
9. Tu JV, Austin PC, Chan BT. Relationship between annual volume of patients treated by admitting physician and mortality after acute myocardial infarction. JAMA. 2001;285:3116-22. [PMID: 11427140]
10. Birkmeyer JD, Stukel TA, Siewers AE, Goodney PP, Wennberg DE, Lucas FL. Surgeon volume and operative mortality in the United States. N Engl J Med. 2003:349:2117-27. [PMID: 14645640]
11. Hartz AJ, Kuhn EM, Pulido J. Prestige ol training programs and experience of bypass surgeons as factors in adjusted patient mortality rates. Med Gate. 1999; 37:93-103. [PMID: 10413397]
12. Rhee SO. Factors determining the quality of physician performance in patient care. Med Care. 1976; 14:733-50. [PMID: 972566]
13. Stolley PD, Becker MH, Lasagna L, McEvilla JD, Sioane LM. The relationship between physician chatacteristics and prescribing appropriateness. Med Care. 1972;10:17-28. [PMID: 5007748]
14. Norcini JJ, Kimball HR, Lipner RS. Certification and specialization: do they matter in the outcome of acute myocardial infarction? Acad Med. 2000;75:1193-8. [PMID: 11112721]
15. Epstein SA, Gonzales JJ, Weinfurt K, Boekeloo B, Yuaji N, Chase G. Ate psychiatrists' characteristics related to how they care for depression in the medically ill? Results from a national case-vignette survey. Psychosomatics. 2001;42: 482-9 [PMID: 118156831
16. Roetzheim RG, Van Durme DJ, Brownlee HJ Jr, Herold AH, Pamies RJ, Woodard L. Compliance with screening mammography. Survey of primary care physicians. J Fla Med Assoc. 1991;78:426-9. [PMID: 1919495]
17. Aubin M, Vézina L, Fortin JP, Bernard PM. Effectiveness of a program to improve hypertension screening in primary care. CMAJ. 1994;150:509-15. [PMID: 8313262]
18. Sanazaro PJ, Worth RM. Measuring clinical performance of individual internists in office and hospital practice. Med Care. 1985;23:1097-114. [PMID: 4033236]
19. Anderson GM, Beers MH, Kerluke K. Auditing prescription practice using explicit criteria and computerized drug benefit claims data. J Eval Clin Pract. 1997;3:283-94. [PMID: 9456428]
20. Epstein SA, Gonzales JJ, Onge JS, Carter-Campbell J, Weinfurt K, Leibole M, et al. Practice patterns in the diagnosis and treatment of anxiety and depression in the medically ill. A survey of psychiatrists. Psychosomatics. 1996;37:356-67. [PMID: 8701014]
21. Adams AS, Soumerai SB, Lomas J, Ross-Degnan D. Evidence of self-report bias in assessing adherence to guidelines. Int J Qual Health Care. 1999; 11:187-92. [PMID: 10435838]
22. Salem-Schatz SR, Avorn J, Soumerai SB. Influence of clinical knowledge, organizational context, and practice style on transfusion decision making. Implications for practice change strategies. JAMA. 1990;264:476-83. [PMID: 2366281]
23. Golden NH, Seigel WM, Fisher M, Schneider M, Quijario E, Suss A, et al. Emergency contraception: pediatricians' knowledge, attitudes, and opinions. Pediatrics. 2001;107:287-92. [PMID: 11158460]
24. Meskauskas JA, Webster GD. The American Board of Internal Medicine recertification examination: process and

results. Ann Intern Med. 1975;82:577-81. [PMID: 1119774]
25. Norcini JJ, Lipner RS, Benson JA Jr, Webster GD. An analysis of the knowledge base of practicing internists as measured by the 1980 recertification examination. Ann Intern Med. 1985; 102:385-9. [PMID: 3970474]
26. Ramsey PG, Carline JD, Inui TS, Larson EB, LoGerfo JP, Norcini JJ, et al. Changes over time in the knowledge base o practicing internists. JAMA. 1991; 266:1103-7. [PMID: 1865543)
27. Cruft GE, Humphreys JW Jr, Hermann RE, Meskauskas JA. Recertification in surgery, 1980. Arch Surg. 1981;116:1093-6. [PMID: 7020647]
28. Gemson DH, Colombotos J, ElinsonJ, Fordyce EJ, Hynes M, Stoneburner R. Acquired immunodeficiency syndrome prevention. Knowledge, attitudes, and practices of primary care physicians. Arch Intern Med. 1991;151:1102-8. [PMID: 2043012]
29. Lewis CE, Freeman HE, Corey CR. AIDS-related competence of California's primary care physicians. Am J Public Health. 1987:77:795-9. [PMID: 3592031]
30. ShapiroJA. General practitioners' attitudes towards AIDS and their perceived information needs. BMJ. 1989:298:1563-6. [PMID: 2503 II 8]
31. Evans CE, Haynes RB, Gilbert JR, Taylor DW, Sackett DL, Johnston M. Educational package on hypertension for primary care physicians. Can Med Assoc J. 1984;130:719-22. [PMID: 6697280]
32. Schroen AT, Detterbeck FC, Crawford R, Rivera MP, Socinski MA. Beliefs among pulmonologisrs and thoracic surgeons in the therapeutic approach to non-small cell lung cancer. Chest. 2000;118:129-37. [PMID: 10893370]
33. Czaja R, McFall SL, Warnecke RB, Ford L, Kaluzny AD. Preferences of community physicians for cancer screening guidelines. Ann Intern Med. 1994: 120:602-8. [PMID: 8117000]
34. Cook RL, Wicsenfeld HC, Ashton MR, Krohn MA, Zamborsky T, Scholle SH. Barriers to screening sexually active adolescent women for chlamydia: a survey of primary care physicians. J Adolesc Health. 2001;28:204-10. [PMID: 11226843]
35. Richards C, Klabunde C, O'Malley M. Physicians' recommendations for colon cancer screening in women. Too much of a good thing? Am J Prev Med. 1998:15:246-9. [PMID: 9791644]
36. Jacques CH, Jones RL, Houts P, Bauer LC, Dwyer KM, Lynch JC, et al. Reported practice behaviors for medical care of patients with diabetes mellitus by primary-care physicians in Pennsylvania. Diabetes Care. 1991:14:712-7. [PMID: 1954806]
37. Kenny SJ, Smith PJ, Goldschmid MG, NewmanJM, Herman WH. Survey of physician practice behaviors related to diabetes mellitus in the U.S. Physician adherence to consensus recommendations. Diabetes Care. 1993:16:1507-10. [PMID 82994401
38. Marrero DG, Moore PS, Fineberg NS, Langefeld CD, Clark CM Jr. The treatment of patients with insulin-requiring diabetes mellitus by primary care physicians. J Community Health. 1991:16:259-67. [PMID: 1955577]
39. Schwärtl/ JS, Lewis CE, Clancy C, Kinosian MS, Radany MH, Koplan JP. Internists' practices in health promotion an disease prevention. A survey. Ann Intern Med. 1991:114:46-53. [PMID: 1983932]
40. Sherman SE, Hershman WY. Exercise counseling: how do general internists do? J Gen Intern Med. 1993:8:243-8. [PMID: 8505682]
41. Zerr DM, Del Beccaro MA, Cummings P. Predictors of physician compliance with a published guideline on

management of febrile infants. Pediatr Infect Dis J. 1999:18:232-8. [PMID: 10093943]
42. Heath K, Hogg RS, Singer J, Schechter MT, O'Shaughnessy MV, Montaner JS. Physician concurrence with primary care guidelines for persons with HIV disease. Int J STD AIDS. 1997:8:609-13. [PMID: 9310219]
43. Skotniski EM, Woodward C, Hutchison B, Abelson J, Brown J, Norman G. HIV testing practices of primary care physicians: an Ontario survey. Can J Public Health. 1996:87:172-5. [PMID: 8771919]
44. Ely JW, Goerdt CJ, Bergus GR, West CP, Dawson JD, Doebbeling BN. I he effect of physician characteristics on compliance with adult preventive care guidelines. Pam Med. 1998:30:34-9. [PMID: 9460614]
45. Rattay KT, Fulton JE, Galuska DA. Weight counseling patterns of U. S. Pediatricians. Obes Res. 2004:12:161-9. [PMID: 14742855]
46. Ford LG, Hunter CP, Diehr P, Frelick RW, Yates J. Effects of patient management guidelines on physician ptactice patterns: the Community Hospital Oncology Program experience. J Clin Oncol. 1987:5:504-11. [PMID: 3819812]
47. Streja DA, Rabkin SW. Factors associated with implementation of preventive care measures in patients with diabetes mellitus. Arch Intern Med. 1999:159: 294-302. [PMID: 9989542]
48. Anis NA, Lee RE, Ellerbeck EF, Nazir N, Greiner KA, Ahluwalia JS. Direct observation of physician counseling on dietary habits and exercise: patient, physician, and office correlates. Prev Med. 2004;38:198-202. [PMID: 14715212]
49. Hulka BS, Kupper LL, Cassel JC. Physician management in primary care. AmJ Public Health. 1976:66:1173-9. [PMID 1008113]
50. Saraiya M, Cookson ST, Tribble P, Silk B, Cass R, Poonja S, et al. Tuberculosis screening among foreign-born persons applying for permanent US residence. AmJ Public Health. 2002:92:826-9. [PMID: 11988454]
51. McFall SL, Warnecke RB, Kaluzny AD, Aidcen M, Ford L. Physician and practice characteristics associated with judgments about breast cancer treatment. Med Care. 1994;32:106-17. [PMID: 8302103]
52. Montaner JS, Hogg RS, Heath KV, Phillips P, Craib KJ, Schechter MT, et al. Heterogeneity of physician agreement with recommended therapeutic guidelines for the management of HIV-associated disease. Antivir Ther. 1996;!: 15766. [PMID: 11322249]
53. Roy-Byrne P, Russo J, Dugdale DC, Lessler D, Cowley D, Katon W. Undertreatmerrt of panic disorder in primary care role of patient and physician characteristics. J Am Board Fam Pract. 2002; 15:443-50. [PMID: 12463289]
54. Hynes DM. The quality of breast cancer care in local communities: implications for health care reform. Med Care. 1994;32:328-40. [PMID: 8139298]
55. Becker MH, Stolley PD, Lasagna L, McEvilla JD, Sloane LM. Characteristics and attitudes of physicians associated with the prescribing of chloramphenicol. HSMHA Health Rep. 1971:86:993-1003. [PMID: 5138286]
56. Ray WA, Federspiel CF, Schaffner W. Prescribing of cbloramphenicol in ambulatory practice. AJI epidemiologic study among Tennessee Medicaid recipients. Ann Intern Med. 1976;84:266-70. [PMID: 769624]
57. Moride Y, Du Fort GG, Monette J, Ducruet T, Boivin JF, Champoux N, et al. Suboptimal duration of antidepressant treatments in the older ambulatory population of Quebec: association with selected physician characteristics. J Am GeriatrSoc. 2002;50:1365-71. [PMID: 12)64992]
58. Beers MH, Fingold SF, Ouslander JG, Reuben DB, Morgenstern H, Beck JC. Characteristics and quality of prescribing by doctors practicing in nursing homes. J Am Geriatr Soc. 1993;41:802-7. [PMID: 8340556]
59. Dhalla IA, Anderson GM, Mamdani MM, Bronskill SE, Sykora K, Rochon PA. Inappropriate prescribing before and

after nursing home admission. J Am GeriatSoc. 2002:50:995-1000. [PMID: 12110057]
60. Geller SE, Burns LR, Brailer DJ. The impact of nonclinical factors on practice variations: the case of hysterectomies. Health Serv Res. 1996;30:729-50. [PMID: 8591927]
61. Payne BC, Lyons TF, Neuhaus E. Relationships of physician characteristics to performance quality and improvement. Health Sen Res. 1984;!9:307-32. IPMID: 6746295]
62. Fehrenbach SN, Budnitz DS, Gazmararian JA, Krumholz HM. Physician characteristics and the initiation of beta-adrenergic blocking agent therapy after acute myocardial infarction in a managed care population. Am J Manag Care. 2001:7:717-23. [PMID: 11464429]
63. Willison DJ, Soumerai SB, Palmer RH. Association of physician and hospital volume with use of aspirin and reperfusion therapy in acute myocardial infarction. Med Care. 2000:38:1092-102. [PMID: 11078050]
64. Beaulieu MD, Biais R, Jacques A, Battista RN, Lebeau R, Brophy J. Are patients suffering from stable angina receiving optima! medical treatment? QJM. 2001:94:301-8. [PMID: 11391028]
65. Blanc PD, Trupin L, Earnest G, San Pedro M, Katz PP, Yelin EH, et al. Effects of physician-related factors on adult asthma cate, health status, and quality of life. AmJ Med. 2003:114:581-7. [PMID: 12753882]
66. O'Neill L, Lanska DJ, Hartz A. Surgeon characteristics associated with mortality and morbidity following carotid endarterectomy. Neurology. 2000:55:77381. [PMID: 10993995]
67. Katon W, Rutter CM, Lin E, Simon G, Von Korff M, Bush T, et al. Aie there detectable differences in quality of care or outcome of depression across primary care providers? Med Care. 2000;38:552-61. [PMID: 10843308]
68. Davidson W, Molloy DW, Bédard M. Physician characteristics and prescribing for elderly people in New Brunswick: relation to patient outcomes. CMAJ. 1995:152:1227-34. [PMID: 7736373]
69. Burns LR, Wholey DR. The effects of patient, hospital, and physician characteristics on length of stay and mortality. Med Care. 1991;29:251-71. [PMID: 1997754]
70. Carthy P, Harvey I, Brawn R, Watkins C. A study of factors associated with cost and variation in prescribing among GPs. Fam Pract. 2000:17:36-41. [PMID: 10673486]
71. Freiman MP. The rate of adoption of new procedures among physicians. The impact of specialty and practice characteristics. Med Care. 1985:23:939-45[PMID: 4021579]
72. Hlatky MA, Cotugno H, O'Connor C, Mark DB, Pryor DB, Califif RM. Adoption of thrombolytic therapy in the management of acute myocardial infarction. AmJ Cardiol. 1988:61:510-4. [PMID: 3125736]
73. Young MJ, Fried LS, Eisenberg J, Hershey J, Williams S. Do cardiologists have higher thresholds for recommending coronary arteriography than family physicians? Health Serv Res. 1987:22:623-35. [PMID: 3692862]
74. Greer AL. The state of the art versus the state of the science. The diffusion of new medical technologies into practice. Int J Technol Assess Health Care. 1988; 4:5-26. [PMID: 10287114]
75. Hall JA, Irish JT, Roter DL, Ehrlich CM, Miller LH. Satisfaction, gender, and communication in medical visits. Med Care. 1994;32:1216-31. [PMID: 7967860]
76. Blanchard CG, Labrecque MS, Ruckdeschel JC, Blanchard EB. Physician behaviors, patient perceptions, and patient characteristics as predictors of satisfaction of hospitalized adult cancer patients. Cancer. 1990:65:186-92. IPM1D): 2293865
77. Jackson JL, Chamberlin K, Kroenke K. Predictors ot patient satisfaction. Soc Sei Med. 2001;52:609-20. [PMID: 11206657]

78. Schmittdiel J, Selby JV, GrumbacK K, Quesenberry CP Jr. Choice of a personal physician and patient satisfaction in a health maintenance organi/ation. JAMA. 1997:278:1596-9. [PMID: 9370506]
79. Grimshaw JM, Shirran L, Thomas R, Mowatt G, Fraser C, Bero L, et al. Changing provider behavior: an overview of systematic reviews of interventions. Med Care. 2001:39:112-45. [PMID: 11583120]

[Author Affiliation]
Niteesh K. Choudhry, MD; Robert H. Fletcher, MD, MSc; and Stephen B. Soumerai, ScD

[Author Affiliation]
From Harvard Medical School and Harvard Pilgrim Health Care, Brigham and Women's Hospital, and the Harvard PhD Program in Health Policy, Boston, Massachusetts.
Grant Support: By the Harvard Pilgrim Health Care Foundation. Dr. Choudhry is also supported by the Harvard Medical School Fellowship in Pharmaceutical Policy Research, a Frank Knox Scholarship from Harvard University, and a Canadia Institutes of Health Research Postdoctoral Fellowship. Dr. Soumerai is an investigator in the HMO Research Network Center for Education and Research in Therapeutics, funded by the U.S. Agency for Healthcare Research and Quality.
Potential Financial Conflicts of Interest: None disclosed.
Requests for Single Reprints: Niteesh K. Choudhry, MD, Department of Ambulatory Care and Prevention, Harvard Medical School, 133 Brookline Avenue, 6th Floor, Boston, MA 02215; e-mail, choudhry@fas .hai-vard.edu.
Current author addresses are available at www.annals.org.