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

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

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 years in practice for some outcomes but no association for others; 2 (3%) reported that performance initially increased with increasing years in practice, 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.


For author affiliations, see end of text.

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.
Clinical Experience and Quality of Health Care

**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, outcome*). 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 used the Fisher exact test to compare the observed frequencies. We conducted all analyses with SAS, version 8.2 (SAS Institute, Inc., Cary, North Carolina).

**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).

---

**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
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, 1.01 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.

**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 \( \beta \)-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.

Hartz and colleagues (11) specifically assessed the association between experience and mortality rates for sur-

---

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

- **Type of Quality Assessed**
  - ■ Studies in which length of time in practice or age was associated with lower performance for all outcomes.
  - □ Studies in which length of time in practice or age was associated with lower performance for some outcomes; no effect was found for other outcomes.
  - □ Studies in which there was a concave relationship between length of time in practice or age and performance.
  - □ Studies in which no association was found between length of time in practice or age and performance.
  - □ Studies in which length of time in practice or age was associated with higher performance for some outcomes; no effect was found for other outcomes.
  - □ Studies in which length of time in practice or age was associated with higher performance for all outcomes.
Table 1. Studies Relating Length of Time in Practice or Physician Age to Knowledge*

<table>
<thead>
<tr>
<th>Study, Year (Reference)</th>
<th>Knowledge Being Assessed</th>
<th>Persons Studied, n</th>
<th>Results</th>
<th>Multivariate Adjustment for Physician Covariates?</th>
<th>Other Comments</th>
<th>Overall Effect?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Salem-Schatz et al., 1990 (22)</td>
<td>Blood product transfusion, general surgeons, orthopedic surgeons, and anesthesiologists</td>
<td>122</td>
<td>Strong inverse association between years in practice and knowledge of transfusion risks and indications (P = 0.0001)</td>
<td>Yes</td>
<td>Knowledge assessment–based medical literature and NIH consensus conference; response rate, 91%</td>
<td>Consistently negative</td>
</tr>
<tr>
<td>Golden et al., 2001 (23)</td>
<td>Emergency contraception</td>
<td>233 pediatricians</td>
<td>Younger physicians and more recent graduates (P = 0.02) were more likely to identify FDA-approved methods of emergency contraception (age categorized as &lt;40 y, 41–50 y, or &gt;50 y)</td>
<td>No</td>
<td>Response rate, 24%</td>
<td>Consistently negative</td>
</tr>
<tr>
<td>Meskauskas and Webster, 1975 (24)</td>
<td>General medical knowledge</td>
<td>3356 internists certified ≥8 y earlier</td>
<td>Inverse relationship between age and ABIM recertification examination scores (age categorized as &lt;40 y, 40–44 y, 45–49 y, 50–54 y, 55–59 y, 60–64 y, or &gt;65 y)</td>
<td>No</td>
<td>Participants self-selected; tests of significance not presented</td>
<td>Consistently negative</td>
</tr>
<tr>
<td>Nocini et al., 1985 (25)</td>
<td>General medical knowledge</td>
<td>1947 internists</td>
<td>Inverse relationship between ABIM recertification examination scores and age (age categorized as &lt;40 y, 40–49 y, 50–59 y, or &gt;60 y)</td>
<td>No</td>
<td>Participants volunteered; tests of significance not presented</td>
<td>Consistently negative</td>
</tr>
<tr>
<td>Ramsey et al., 1991 (26)</td>
<td>General medical knowledge</td>
<td>289 internists certified 5 to 15 y earlier</td>
<td>Significant inverse correlation (r = –0.3) between score on ABIM examination questions and years since certification</td>
<td>Yes</td>
<td>Participants partially self-selected, but sample was representative of population</td>
<td>Consistently negative</td>
</tr>
<tr>
<td>Cruft et al., 1981 (27)</td>
<td>General surgical knowledge</td>
<td>478 surgeons certified ≥7 y earlier</td>
<td>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)</td>
<td>No</td>
<td>Participants self-selected; tests of significance not presented</td>
<td>Consistently negative</td>
</tr>
<tr>
<td>Gemerson et al., 1991 (28)†</td>
<td>HIV</td>
<td>473 internists, family practitioners, general practitioners, and obstetrician–gynecologists</td>
<td>Younger physicians had significantly more knowledge about AIDS (P = 0.01)</td>
<td>Yes</td>
<td>Knowledge assessed by using questions from National Center for Health Statistics survey and others devised by investigators; response rate, 63%</td>
<td>Consistently negative</td>
</tr>
<tr>
<td>Lewis et al., 1987 (29)†</td>
<td>HIV</td>
<td>1000 general practitioners, family physicians, and general internists</td>
<td>Younger physicians had greater AIDS-related knowledge (consistent across variables assessed, although P values not reported)</td>
<td>Unclear</td>
<td>Measures of “competence” used were defined by group of expert clinicians at UCLA; response rate, 60%</td>
<td>Consistently negative</td>
</tr>
<tr>
<td>Shapiro, 1983 (30)</td>
<td>HIV</td>
<td>1271 general practitioners</td>
<td>Knowledge of HIV and AIDS decreased as years since graduation increased (P = 0.0008)</td>
<td>Yes</td>
<td>Knowledge assessed with 6 questions designed by author; response rate, 70%</td>
<td>Consistently negative</td>
</tr>
<tr>
<td>Evans et al., 1984 (31)</td>
<td>Hypertension</td>
<td>56 family physicians</td>
<td>Highly significant inverse correlation between test scores and years since graduation (r = –0.55; P &lt; 0.001)</td>
<td>No</td>
<td>Questionnaire validated to discriminate among physicians of different levels of training and specialty; response rate, 78%</td>
<td>Consistently negative</td>
</tr>
<tr>
<td>Ayasian et al., 1994 (7)</td>
<td>MI</td>
<td>1211 cardiologists, internists, and family practitioners</td>
<td>Physicians &lt; 40 y of age had greater knowledge of evidence-based therapies (P &lt; 0.05)</td>
<td>Yes</td>
<td>All physicians had served as the attending for at least 1 patient with MI within the preceding 3 mo; response rate, 61%</td>
<td>Consistently negative</td>
</tr>
<tr>
<td>Schroen et al., 2000 (32)</td>
<td>Non-small-cell lung cancer</td>
<td>1010 pulmonologists and thoracic surgeons</td>
<td>Physicians trained before 1980 were less likely to underestimate survival (P &lt; 0.001) and less likely to believe in value of chemotherapy in situations that have been well-established</td>
<td>No</td>
<td>Response rate, approximately 50%</td>
<td>Consistently negative</td>
</tr>
</tbody>
</table>

* 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.
### 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>
<thead>
<tr>
<th>Study, Year (Reference)</th>
<th>Disease or Condition</th>
<th>Physician Group Studied</th>
<th>Sample Size, n</th>
<th>Results</th>
<th>Multivariate Adjustment for Patient Covariates?</th>
<th>Multivariate Adjustment for Physician Covariates?</th>
<th>Other Comments</th>
<th>Overall Effect†</th>
</tr>
</thead>
<tbody>
<tr>
<td>Czaja et al., 1994 (33)</td>
<td>Cancer screening</td>
<td>Family physicians, internists, general surgeons, and gynecologists</td>
<td>3436 physicians</td>
<td>Physicians who graduated &gt;20 y ago less likely to adhere to screening practices (OR: 0.62–0.72; P &lt; 0.05)</td>
<td>No</td>
<td>Yes</td>
<td>Only considered adherence with guidelines for use of interventions unanimously endorsed by multiple organizations; response rate, 67%</td>
<td>Consistently negative</td>
</tr>
<tr>
<td>Cook et al., 2001 (34)</td>
<td>Chlamydia screening</td>
<td>Family physicians, internists, gynecologists, and pediatricians</td>
<td>1600 physicians</td>
<td>No effect of length of time in practice and likelihood of screening (&lt;10 y or &gt;10 y in practice)</td>
<td>No</td>
<td>Yes</td>
<td>Guidelines established by CDC and USPSTF</td>
<td>No effect</td>
</tr>
<tr>
<td>Richards et al., 1998 (35)</td>
<td>Colon cancer screening for women</td>
<td>Primary care providers</td>
<td>508 physicians</td>
<td>Older physicians more likely to recommend screening contrary to national guidelines (OR, 3.42–3.79)</td>
<td>No</td>
<td>Yes</td>
<td>Response rate, 42%</td>
<td>Consistently negative</td>
</tr>
<tr>
<td>Epstein et al., 2001 (15)‡</td>
<td>Depression</td>
<td>Psychiatrists</td>
<td>278 physicians</td>
<td>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)</td>
<td>No</td>
<td>Yes</td>
<td>Appropriateness defined by consensus agreement of 4 national experts; analyses adjusted for medical comorbidity but not severity of depression; response rate, 53%</td>
<td>Consistently negative</td>
</tr>
<tr>
<td>Jacques et al., 1991 (36)</td>
<td>Diabetes</td>
<td>General practitioners, family physicians, and general internists</td>
<td>610 physicians</td>
<td>Physicians who had graduated more recently more likely to appropriately use glucose self-monitoring, hemoglobin A$_1c$ measurements, ophthalmology examinations (P &lt; 0.001); no effect was observed for blood pressure and weight assessment, foot examination, glycemic education; year of graduation categorized in 4 groups</td>
<td>No</td>
<td>Yes</td>
<td>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)</td>
<td>Partially negative</td>
</tr>
<tr>
<td>Kenny et al., 1993 (37)</td>
<td>Diabetes</td>
<td>Primary care physicians</td>
<td>1434 physicians</td>
<td>Younger physicians more likely to adhere to preventive care guidelines for 6 of 8 procedures surveyed</td>
<td>No</td>
<td>Yes</td>
<td>Effect estimates not reported but results based on logistic regression; guidelines established by American Diabetes Association</td>
<td>Partially negative</td>
</tr>
<tr>
<td>Marrero et al., 1991 (38)</td>
<td>Diabetes</td>
<td>Primary care physicians</td>
<td>212 physicians</td>
<td>Younger physicians more likely to obtain a hemoglobin A$_1c$ measurement (OR for every 10-y change in graduation date, 1.53; P = 0.0017); no relationship for use of glucose self-monitoring</td>
<td>No</td>
<td>No</td>
<td>Guidelines established by American Diabetes Association; response rate, 31%</td>
<td>Partially negative</td>
</tr>
<tr>
<td>Schwartz et al., 1991 (39)</td>
<td>Diabetes prevention and health promotion</td>
<td>Members and Fellows of ACP</td>
<td>1349 physicians</td>
<td>Appropriate use of health promotion and disease prevention practices decreased with increasing age (P value not presented)</td>
<td>No</td>
<td>Yes</td>
<td>Guidelines endorsed by several national agencies; response rate, 75%</td>
<td>Consistently negative</td>
</tr>
<tr>
<td>Sherman and Hershman, 1993 (40)</td>
<td>Exercise counseling</td>
<td>Primary care physicians</td>
<td>422 physicians</td>
<td>Physicians &gt;40 y of age more likely to counsel patients on exercise (OR, 3.08 [CI, 1.33–7.15])</td>
<td>No</td>
<td>Yes</td>
<td>Justification for exercise counseling based on research evidence; response rate, 61%</td>
<td>Consistently positive</td>
</tr>
<tr>
<td>Zerr et al., 1999 (41)</td>
<td>Fever in infants</td>
<td>Pediatrists and emergency department and family physicians</td>
<td>474 physicians</td>
<td>Physicians who had graduated longer ago less likely to adhere to guidelines (OR, 0.93 [CI, 0.91–0.96], per year since graduation)</td>
<td>No</td>
<td>Yes</td>
<td>Physicians provided with guidelines; adherence assessed by using clinical scenarios that presented children of different ages; response rate, 36%</td>
<td>Consistently negative</td>
</tr>
<tr>
<td>Gerson et al., 1991 (28)</td>
<td>HIV</td>
<td>Internists, family doctors, general practitioners, and obstetrician–gynecologists</td>
<td>473 physicians</td>
<td>Inverse relationship between years since graduation and adherence to New York State Department of Health AIDS prevention recommendations (P &lt; 0.01)</td>
<td>No</td>
<td>Yes</td>
<td>Study also assessed knowledge (results presented separately); response rate, 63%</td>
<td>Consistently negative</td>
</tr>
<tr>
<td>Heath et al., 1997 (42)</td>
<td>HIV</td>
<td>Family physicians and specialists who treat HIV</td>
<td>868 physicians</td>
<td>Inverse relationship between physician age and use of appropriate preventive care strategies (P &lt; 0.001–0.004)</td>
<td>No</td>
<td>Yes</td>
<td>Response rate, 38.2%–50%; guidelines were issued by provincial agency</td>
<td>Consistently negative</td>
</tr>
</tbody>
</table>

Continued on following page
Table 2—Continued

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

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

<table>
<thead>
<tr>
<th>Study, Year (Reference)</th>
<th>Disease or Condition</th>
<th>Physician Group Studied</th>
<th>Sample Size, n</th>
<th>Results</th>
<th>Multivariate Adjustment for Patient Covariates?</th>
<th>Multivariate Adjustment for Physician Covariates?</th>
<th>Other Comments</th>
<th>Overall Effect†</th>
</tr>
</thead>
<tbody>
<tr>
<td>Self-reported practice (surveys or interviews)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>McFall et al., 1994 (51)</td>
<td>Breast cancer</td>
<td>Family physicians, internists, gynecologists, and general surgeons</td>
<td>1460 physicians</td>
<td>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</td>
<td>No</td>
<td>Yes</td>
<td>Physicians were not oncologists but reported participating in decision making about treatment and referral; response rate, 71%</td>
<td>Partially negative</td>
</tr>
<tr>
<td>Epstein et al., 2001 (15)</td>
<td>Depression</td>
<td>Psychiatrists</td>
<td>278 physicians</td>
<td>Physician age did not influence the appropriate prescribing of an antidepressant</td>
<td>Limited</td>
<td>Yes</td>
<td>Appropriateness defined by consensus of 4 national experts; analyses adjusted for medical comorbidity but not severity of depression; response rate, 53%</td>
<td>No effect</td>
</tr>
<tr>
<td>Epstein et al., 1996 (20)</td>
<td>Depression and anxiety</td>
<td>Psychiatrists</td>
<td>38 physicians</td>
<td>“Accuracy score” (reflecting agreement with expert consensus) decreased as a function of years in practice (P &lt; 0.01)</td>
<td>No</td>
<td>Yes</td>
<td>Appropriateness established by expert consensus; response rate, 19%</td>
<td>Consistently negative</td>
</tr>
<tr>
<td>Montaner et al., 1996 (52)</td>
<td>HIV</td>
<td>Physicians in British Columbia</td>
<td>463 physicians</td>
<td>Physicians &lt; 45 y of age significantly more likely to appropriately use antiretroviral therapy (P = 0.004); no age effect observed for other management areas</td>
<td>No</td>
<td>Yes</td>
<td>Appropriateness assessed by using provincial guidelines; overall response rate, 14%</td>
<td>Partially negative</td>
</tr>
<tr>
<td>Roy-Byrne et al., 2002 (53)</td>
<td>Panic disorder</td>
<td>Primary care physicians</td>
<td>37 physicians (58 patients)</td>
<td>Length of time in practice did not predict appropriate prescribing</td>
<td>Yes</td>
<td>Yes</td>
<td>Based on patient reports (all enrolled in clinical trial); appropriateness based on previously published algorithm</td>
<td>No effect</td>
</tr>
<tr>
<td>Stolley et al., 1972 (13)</td>
<td>Prescribing of 5 specific drugs (ritalin, equagesic, chloromycetin, vitamin B12, and oral contraceptives)</td>
<td>Primary care physicians</td>
<td>37 physicians</td>
<td>Appropriateness decreased as years in practice increased (P &lt; 0.01)</td>
<td>No</td>
<td>No</td>
<td>Appropriateness assessed by at least 13 experts who rated any given drug, and the total panel consisted of 33 individuals; response rate, 84%</td>
<td>Consistently negative</td>
</tr>
<tr>
<td>Hynes, 1994 (54)</td>
<td>Breast cancer</td>
<td>Physican treating breast cancer</td>
<td>Not reported (3972 patients)</td>
<td>Appropriateness of surgical care increased with increasing years in practice but decreased after 14 y of experience (P &lt; 0.01); physicians in practice for more years were less likely to provide postmastectomy rehabilitation therapy</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
<td>Concave</td>
</tr>
<tr>
<td>Becker et al., 1971 (55)</td>
<td>Chloramphenicol use</td>
<td>Primary care physicians</td>
<td>37 physicians</td>
<td>Likelihood of prescribing chloramphenicol increased as years since graduation increased (P &lt; 0.01)</td>
<td>No</td>
<td>Yes</td>
<td>All prescriptions of chloramphenicol were judged to be inappropriate given limited indications</td>
<td>Consistently negative</td>
</tr>
<tr>
<td>Ray et al., 1976 (56)</td>
<td>Chloramphenicol use</td>
<td>Physicians in Tennessee caring for Medicaid patients</td>
<td>3409 physicians</td>
<td>Number of years since medical school graduation did not predict use of chloramphenicol</td>
<td>No</td>
<td>Yes</td>
<td>All prescriptions of chloramphenicol were judged to be inappropriate given limited indications</td>
<td>No effect</td>
</tr>
<tr>
<td>Moride et al., 2002 (57)</td>
<td>Depression</td>
<td>General practitioners and psychiatrists</td>
<td>1527 physicians</td>
<td>Graduation from medical school before 1970 associated with increased odds of suboptimal treatment duration (OR, 1.12 [95% CI, 1.01–1.24])</td>
<td>Yes</td>
<td>Yes</td>
<td>Patient covariates include age, sex, and health status but not illness severity</td>
<td>Consistently negative</td>
</tr>
</tbody>
</table>

Continued on following page
“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 and 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

### Table 3—Continued

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

*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.
those studies that used objective measures (71% vs. 62%), these differences were not statistically significant ($P > 0.2$).

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 physicians 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

---

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

<table>
<thead>
<tr>
<th>Study, Year (Reference)</th>
<th>Disease or Condition</th>
<th>Physician Group Studied</th>
<th>Sample Size, n</th>
<th>Results</th>
<th>Multivariate Adjustment for Patient Covariates?</th>
<th>Multivariate Adjustment for Physician Covariates?</th>
<th>Other Comments</th>
<th>Overall Effect*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nocini et al., 2000 (14)</td>
<td>Acute myocardial infarction</td>
<td>Family practitioners, interns, and cardiologists</td>
<td>4546 physicians</td>
<td>Mortality for patients admitted with acute myocardial infarction increased by 0.5% for every year since physician graduated from medical school ($P = 0.05$)</td>
<td>Yes</td>
<td>Yes</td>
<td>Also corrected for hospital factors (e.g., access to advanced cardiac care)</td>
<td>Consistently negative</td>
</tr>
<tr>
<td>Blanc et al., 2003 (65)</td>
<td>Asthma</td>
<td>Physicians treating asthma</td>
<td>147 physicians (317 patients)</td>
<td>Number of years since medical school graduation not related to patients' self-reported health status or asthma-specific quality of life</td>
<td>Yes</td>
<td>Yes</td>
<td>Surveyed patients and physicians separately and then linked results using hierarchical regression models; sample included very few younger physicians</td>
<td>No effect</td>
</tr>
<tr>
<td>O’Neill et al., 2000 (66)</td>
<td>Carotid endarterectomy</td>
<td>Surgeons</td>
<td>507 physicians (12,725 patients)</td>
<td>Mortality of patients undergoing endarterectomy increased with years since licensure ($P &lt; 0.0001$); no relationship between length of time in practice and combined “bad outcome” (death or morbidity)</td>
<td>Yes</td>
<td>Yes</td>
<td>Data for surgeon age available for 440 physicians; years since licensure was strongest predictor of mortality</td>
<td>Partially negative</td>
</tr>
<tr>
<td>Hartz et al., 1999 (11)</td>
<td>Coronary bypass surgery</td>
<td>Surgeons</td>
<td>275 physicians (83,457 patients)</td>
<td>More years in practice significantly associated with higher mortality ratios ($C = 0.22$, $P &lt; 0.001$)</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
<td>Consistently negative</td>
</tr>
<tr>
<td>Katon et al., 2000 (67)</td>
<td>Depression</td>
<td>Primary care physicians</td>
<td>63 physicians (1,599 patients)</td>
<td>Physician age not related to patients’ having persistent or residual depressive symptoms after initiation of antidepressant medication or other quality-of-care measures</td>
<td>Yes</td>
<td>Yes</td>
<td>Study did not find any significant physician characteristics to explain variability in quality of care</td>
<td>No effect</td>
</tr>
<tr>
<td>Davidson et al., 1995 (68)</td>
<td>Drug prescribing in elderly patients</td>
<td>General practitioners</td>
<td>366 physicians</td>
<td>No age effect on mortality or hip fracture rate</td>
<td>Limited</td>
<td>No</td>
<td>Adjusted for patient age by using analysis of covariance</td>
<td>No effect</td>
</tr>
<tr>
<td>Burns and Wholey, 1991 (69)</td>
<td>Several (11 medical and 5 surgical conditions)</td>
<td>Attending physicians for hospitalized patients</td>
<td>54,571 discharges</td>
<td>Physicians in practice for more years had significantly longer lengths of stay after adjustment for patient comorbid condition for 9 of 16 diagnoses evaluated ($P &lt; 0.05$); years in practice did not predict length of stay for other diagnoses or for mortality</td>
<td>Yes</td>
<td>Yes</td>
<td>Results adjusted for multiple covariates, suggesting increased length of stay may be unnecessary</td>
<td>Partially negative</td>
</tr>
</tbody>
</table>

* “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.
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.

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 Postdoctoral Fellowship. Dr. Soumerai is an investigator in the HMO Research Network Center for Education and Research in Therapeutics, funded by a Frank Knox Scholarship from Harvard University, and a Canadian Institutes of Health Research Postdoctoral Fellowship in Pharmaceutical Policy Research, a Canadian 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.harvard.edu.

Current author addresses are available at www.annals.org.

References


www.annals.org
55. 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]


Current Author Addresses: Drs. Choudhry, Fletcher, and Soumerai:
Department of Ambulatory Care and Prevention, Harvard Medical
School, 133 Brookline Avenue, 6th Floor, Boston, MA 02215.