Comparison of Physician Workforce Estimates and Supply Projections

Douglas O. Staiger, PhD
David I. Auerbach, PhD
Peter I. Buerhaus, PhD, RN

RECENT PROJECTIONS INDICATE that the supply of US physicians may soon decrease below requirements, with some projecting a shortfall as high as 200,000 by 2020. Although debate over potential shortages has focused largely on the number and type of physicians needed in the future, concerns have also been raised about data used in physician supply estimates and projections. The American Medical Association Physician Masterfile (Masterfile) data are believed to overestimate the number of active physicians at older ages. The overestimation has been attributed to delays in updating the Masterfile data when a physician retires or experiences a change in status, such as specialty, location, or immigration. The US Health Resources and Services Administration (HRSA) and the Council on Graduate Medical Education acknowledge that these delays may lead to overestimates of both current and future physician supply. To account for this concern, recent workforce projections have applied various adjustments when using the Masterfile data.

We conducted parallel retrospective cohort analyses of employment trends of physicians over their lifetimes using the Masterfile data and the US Census Bureau Current Population Survey (CPS), a data source used extensively by the US Department of Labor to estimate current trends in employment, and previously used to estimate employment trends for registered nurses. In addition, physician supply through 2040 was projected using both data sources. These comparisons were performed to provide empirical estimates of the differences between the Masterfile data and CPS.

Objective To compare physician workforce estimates and supply projections using the American Medical Association Physician Masterfile (Masterfile) data with estimates and projections using data from the US Census Bureau Current Population Survey (CPS).

Design, Setting, and Participants Parallel retrospective cohort analyses of employment trends of the number of active physicians by age and sex using annual data from the Masterfile and the CPS between 1979 and 2008. Recent workforce trends were used to project future physician supply by age.

Main Outcome Measure Annual number of physicians working at least 20 hours per week in 10-year age categories.

Results In an average year in the sample period, the CPS estimated 67,000 (10%) fewer active physicians than did the Masterfile (95% confidence interval [CI], 57,000-78,000; P < .001), almost entirely due to fewer active physicians aged 55 years or older. The CPS estimated more young physicians (ages 25-34 years) than did the Masterfile, with the difference increasing to an average of 17,000 (12%) during the final 15 years (95% CI, 13,000-22,000; P < .001). The CPS estimates of more young physicians were consistent with historical growth observed in the number of first-year residents, and the CPS estimates of fewer older physicians were consistent with lower Medicare billing by older physicians. Projections based on both the CPS and the Masterfile data indicate that the number of active physicians will increase by approximately 20% between 2005 and 2020. However, projections for 2020 using CPS data estimate nearly 100,000 (9%) fewer active physicians than projections using the Masterfile data (957,000 vs 1,050,000), and estimate that a smaller proportion of active physicians will be 65 years or older (9% vs 18%). The increasing proportion of female physicians had little effect on physician supply projections because, unlike male physicians, female physicians were found to maintain their work activity after age 55 years.

Conclusion Compared with the Masterfile data, estimates using the CPS data found more young physicians entering the workforce and fewer older physicians remaining active, resulting in estimates of a smaller and younger physician workforce now and in the future.

JAMA. 2009;302(15):1674-1680 www.jama.com

For editorial comment see p 1701.
PHYSICIAN WORKFORCE ESTIMATES AND SUPPLY PROJECTIONS

Masterfile and the CPS-based estimates of the size of the physician workforce both among older and younger physicians and by sex.

METHODS
Data
The Masterfile is continuously updated and contains demographic, educational, and current professional activity on more than 940,000 medical and osteopathic physicians, including AMA members and nonmembers. The Masterfile uses various sources to identify all physicians in the United States upon entry into medical school or into the United States, and monitors physician activity using a rotating census in which approximately one-third of physicians are surveyed each year. At the end of each year, the current data are used to count the number of active physicians. The main strength of the Masterfile is that these data provide a census of physicians by specialty and practice. The main limitation is a lag in recording changes in physician status, such as new licensure, change in specialty, and retirement.

The Masterfile data were obtained from published counts of active medical physicians working at least 20 hours per week by age and sex reported in the 1982 through 2008 editions of the annual publication Physician Characteristics and Distribution in the US. These counts included residents, physicians not directly involved in patient care, and physicians who were not classified. Twenty-four years of data were available for 5 age groups (<35 years, 35-44 years, 45-54 years, 55-64 years, and ≥65 years) between 1980 and 2007, excluding 1984, 1987, 1991, and 1994, when data were not published. Counts of osteopathic physicians by the same age groups and by sex were obtained from the American Osteopathic Association for available years (1984, 1994, 1998, 2001, 2004, and 2007) (Steven Andes, PhD, American Osteopathic Association, written communication, 2009), with intervening years estimated assuming constant growth rates between available years. Osteopathic physicians accounted for less than 6% of physicians throughout our sample period.

The CPS, a household-based survey administered monthly by the US Census Bureau, selects a nationally representative sample of the noninstitutionalized civilian population, including more than 300,000 individuals each year. Current Population Survey data are used to estimate current trends in unemployment, employment, and earnings; have survey response rates of more than 90%; and are validated by the US Department of Labor.

Approximately 1000 individuals identify their occupation as “physician [or] surgeon” each year. Although CPS data introduce sampling error into estimates of the physician workforce, its main strength is the avoidance of systematic biases such as reporting lags. The main limitation is the exclusion of physicians on active military duty (<2% of physicians) and the possible exclusion of licensed physicians who do not report their occupation as “physician [or] surgeon.” Approximately 5% of active physicians in the Masterfile are involved in “other professional activities” (administration, teaching, and research) of which some proportion may not be captured by the CPS. The CPS sample provided data for 32,019 physicians who reported working at least 20 hours in the week of the survey between 1979 and 2008. Annual estimates were constructed using sampling weights provided by the CPS.

Two additional data sources were used to construct benchmarks of physician activity for comparison to the Masterfile and CPS data. First, to capture trends in the number of young physicians, we used the numbers of first-year medical residents obtained from the 1982 through 2008 Medical Education issues of JAMA and supplemented by additional data on residents in American Osteopathic Association internships. Second, to measure the decrease in activity at older ages, we used Medicare Part B physician relative value units from 2002-2004 obtained from the Centers for Medicare & Medicaid Services 20% sample of Medicare beneficiaries, linked to data on physician age in 2004 from the Masterfile.

Statistical Analysis
Trend Analysis. The CPS data were used to estimate the number of US active physicians in 10-year age categories corresponding to data reported by the Masterfile (<35 years, 35-44 years, 45-54 years, 55-64 years, and ≥65 years). Sampling error for CPS estimates was calculated by using formulas provided by the US Bureau of Labor Statistics. The Masterfile data are based on a census and have no associated sampling error. Confidence intervals (CIs) for the mean difference between the CPS-derived and Masterfile-derived estimates were constructed by using standard 2-sample t statistics, based on 24 annual observations of the difference between the 2 estimates. All P values were based on 2-tailed tests of significance.

Statistical Model. The analysis relied on a statistical model commonly used by demographers and economists; a detailed description is presented elsewhere. Briefly, the model decomposes observed changes in the size and age of the physician workforce over time into the product of 2 components: cohort and age effects. The cohort effect refers to the number of individuals born in any given 10-year period (eg, 1951-1960) who will be active physicians when the cohort reaches the ages of 45 to 54 years (eg, in 2005), and captures the size of a given cohort of physicians in their peak working years. The age effect refers to the relative propensity of physicians to work when the cohort reaches other age ranges (<35 years, 35-44 years, 55-64 years, and ≥65 years) relative to ages 45 to 54 years, and captures lifecycle patterns such as retirement and the tendency of female physicians to work less during childbearing years. Thus, the number of physicians working as each cohort reaches a given age range is the
product of the size of the cohort and the propensity to be working at that age.

Estimation. Analysis of variance was used to estimate age and cohort effects separately using the CPS and Masterfile data, and separately for men and women. Because the Masterfile data were only available in 5 age categories, the equation estimates 5 distinct age effects in addition to 69 distinct cohort effects, 1 for each 10-year span beginning 1905 through 1974. Cohort effects for older physicians were derived from employment data collected only late in the career, assuming that these physicians’ retirement patterns (ie, age effects) were the same as those for more recent cohorts. Estimates of age effects changed very little between the first and second half of our sample, suggesting that stable age effects are a reasonable assumption. The number of observations in the model estimated using CPS data was 150 (5 age categories spanning 30 years of data [1979-2008]). Because of missing years in the Masterfile data, the analysis of the Masterfile data used 120 observations. All statistical analyses were performed by using Stata version 10.1 (Stata Corp, College Station, Texas).

Projections. Estimates of age and cohort effects from the CPS and Masterfile data were used to project the numbers of active physicians for the years between 2020 and 2040. We assumed that age effects in future years will be the same as in past years (ie, the propensity of a physician aged 35-44 years to be active relative to a physician aged 45-54 years will not differ in the future) and that the cohort effect for future cohorts (entering the workforce after 2008) will equal the average of the 5 most recent cohorts observed (the 1970-1979 through 1974-1983 cohorts). These assumptions mean the cohorts already in the workforce would follow the same lifecycle pattern as that observed in recent cohorts and that the size of new cohorts entering the workforce would remain constant at recently observed levels.

These projections were compared with recent projections of the number of active physicians from HRSA to determine whether differences in underlying data, rather than the use of any particular projection methodology, drive the differences between the projections. HRSA used the Masterfile data from 2000 to estimate the initial size of the workforce and to estimate retirement rates, and used data on the numbers of physicians in Graduate Medical Education to estimate physician entry. The HRSA baseline projection model assumed that retirement rates and physician entry would remain at recent levels. Because our model makes similar assumptions, we expected it to yield projections similar to HRSA when applied to the Masterfile data through 2000.

RESULTS

Trends in the Number of Active Physicians

Analysis of the CPS and Masterfile data indicate that the number of active physicians of all ages approximately doubled during the sample period (Figure 1). In an average year, the CPS estimated 67,000 (10%) fewer active physicians than did the Masterfile during the sample period (95% CI, 57,000-78,000; P < .001). The difference between the 2 estimates was relatively stable over time and did not change significantly between the first (1979-1993) and second halves (1994-2008) of the sample (P = .38).

Figure 1 also shows the differences in the estimates by age group. Estimates from the Masterfile and CPS data were similar for physicians between the ages of 35 and 54 years, but differ markedly at both younger and older ages. Lower estimates of active physicians in the CPS were almost entirely accounted for by older physicians. On average, the CPS estimated 22,000 (20%) fewer active physicians per year aged 55 to 64 years than did the Masterfile during the sample period (95% CI, 17,000-22,000; P < .001), and estimated 35,000 (51%) fewer active physicians per year 65 years or older rela-
The CPS estimates indicating fewer active physicians at older ages were consistent with data on physician billing patterns available from Medicare Part B for the years 2002-2004, which indicate a sharp decrease in billing activity among older physicians. Total billing by physicians aged 65 years or older was 12% of total billing by physicians aged 45 to 54 years during 2002-2004. Similarly, in the CPS, the number of active physicians aged 65 years or older was 15% of the total number of active physicians aged 45 to 54 years during 2002-2004. However, in the Masterfile data, the number of active physicians aged 65 years or older was 39% of the total number of active physicians aged 45 to 54 years during 2002-2004, consistent with an overestimation of the activity rate of physicians at older ages.

With respect to younger physicians, the CPS on average estimated 9000 (6%) more active physicians aged 25 to 34 years than did the Masterfile during the sample period (95% CI, 29 000-40 000; P < .001). The CPS estimates differed significantly between the data sources implying a more rapid decrease in activity at older ages. For example, relative to the number of physicians aged 25 to 34 years, data from the Masterfile estimated that 86% as many physicians will remain active between the ages of 65 to 74 years (95% CI, 83%-89%), although the CPS data estimated that only 44% as many physicians will remain active (95% CI, 41%-48%).

The age effects trace out different patterns by sex, likely reflecting the effects of childbearing female physicians. Using CPS data, compared with the peak work years of 45 to 54 years, female physicians are 59% as active at 25 to 34 years and 98% as active at 55 to 64 years. In contrast, relative to their own peak activity rates, which also come at ages 45 to 54 years, male physicians are 68% as active at 25 to 34 years, but only 83% as active at 55 to 64 years. Estimates from the Masterfile data display a similar pattern. Jointly, the estimated age effects differed significantly between men and women in both the CPS (P < .001) and Masterfile (P = .03) data, although individually the CIs on the age effect are wide in the CPS data.

### Table 1. Expected Number of Active Physicians When a Cohort Reaches Each Age Range as a Proportion of the Number of Active Physicians When Each Cohort Is Aged 45 to 54 Years

<table>
<thead>
<tr>
<th>Age Range, y</th>
<th>Masterfile Data Used to Estimate Age Effects</th>
<th>CPS Data Used to Estimate Age Effects</th>
</tr>
</thead>
<tbody>
<tr>
<td>Separately by Sex</td>
<td>Pooled</td>
<td>Men</td>
</tr>
<tr>
<td>25-34</td>
<td>0.62 (0.60-0.65)</td>
<td>0.65 (0.62-0.67)</td>
</tr>
<tr>
<td>35-44</td>
<td>0.97 (0.94-1.00)</td>
<td>0.98 (0.95-1.01)</td>
</tr>
<tr>
<td>55-64</td>
<td>0.95 (0.92-0.98)</td>
<td>0.94 (0.91-0.97)</td>
</tr>
<tr>
<td>65+</td>
<td>0.86 (0.83-0.89)</td>
<td>0.84 (0.81-0.87)</td>
</tr>
</tbody>
</table>

### Projection of Future Physician Supply

TABLE 2 compares the projected number and percentage change of active physicians between 2005 and 2040.
Projections were constructed by using the CPS data through 2008 and the Masterfile data through 2007. Both projections allowed age and cohort effects to differ by physician sex. Compared with projections that assumed the same age effects for male and female physicians, allowing age effects to differ by sex yields a slightly larger growth in the projected workforce, reflecting that large cohorts of female physicians will soon be reaching older ages, at which time they will maintain their activity rates unlike men who have tended to work relatively less at those ages. The HRSA baseline and high-growth physician requirements scenarios are provided as a benchmark. Supply projections based on either the CPS or Masterfile data yield similar workforce growth rates between 2005 and 2020 (21.4% for the CPS data and 23.1% for the Masterfile data), reflecting 2 offsetting differences. Compared with the Masterfile data, the CPS data estimates that more physicians will exit the workforce as an increasing number of physicians reach 55 years or older, while estimating that more young physicians have been entering the workforce in recent years. Both supply projections yield growth rates close to the growth in physician requirements projected by the HRSA baseline scenario (21.7%).

However, CPS-based projections suggest that the number of active physicians will be below HRSA’s baseline requirements by 19,000 physicians in 2020, and the Masterfile-based projections suggest that the number of active physicians will exceed HRSA’s baseline requirements by 74,000 physicians—a difference between the projections of nearly 100,000 physicians (9%). This gap primarily reflects differences in how many physicians were estimated to be active in 2005, rather than differences in projected growth rates. Both projections are below the HRSA high-growth scenario, with the CPS-based projection suggesting a deficit of 155,000 physicians and the Masterfile-based projection suggesting a deficit of 62,000 physicians.

Because the CPS data result in estimates of both larger cohorts of young physicians and greater decreases in activity at older ages relative to the Masterfile data, CPS-based projections result in a younger distribution of active physicians in 2020 (FIGURE 3). The CPS-based projection indicates that 71% of active physicians will be younger than 55 years and only 9% will be older than 65 years, whereas the Masterfile-based projection indicates that 61% of active physicians will be younger than 55 years and 18% will be older than 65 years.

These projections were compared with the Masterfile-based projections of physician supply constructed by HRSA (which used somewhat different meth-

---

**Table 2. Projection of Total Number of Active Physicians Through 2040 Based on the Masterfile and CPS Data vs HRSA Estimates of Physician Requirements Through 2020**

<table>
<thead>
<tr>
<th></th>
<th>Projected No. of Active Physicians (in 1000s)</th>
<th>% Change</th>
<th>2005-2020</th>
<th>2005-2030</th>
<th>2005-2040</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data used to form projections</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CPS</td>
<td>788</td>
<td>957</td>
<td>1013</td>
<td>1032</td>
<td>21.4</td>
</tr>
<tr>
<td>Masterfile</td>
<td>853</td>
<td>1050</td>
<td>1103</td>
<td>1133</td>
<td>23.1</td>
</tr>
<tr>
<td>HRSA requirements</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline</td>
<td>802</td>
<td>976</td>
<td>NA</td>
<td>NA</td>
<td>21.7</td>
</tr>
<tr>
<td>High growth</td>
<td>802</td>
<td>1112</td>
<td>NA</td>
<td>NA</td>
<td>38.7</td>
</tr>
</tbody>
</table>

Abbreviations: CPS, US Census Bureau Current Population Survey; Masterfile, American Medical Association Physician Masterfile. Estimates are based on the model as described in the “Statistical Analysis” section of the text. Each line plots estimates based on data from a different sample (men, women, and pooled) and data source (Masterfile and CPS). Standard errors on the Masterfile data estimates for each birth cohort are 2% to 5%; on the CPS data, 7% to 12%, except for female cohorts (22%-38%). Estimates for more recent cohorts are inferred from employment at ages 25 to 34 years (when CPS estimates had greater employment than the Masterfile data, CPS-based projections result in a younger distribution of active physicians in 2020 (FIGURE 3). The CPS-based projection indicates that 71% of active physicians will be younger than 55 years and only 9% will be older than 65 years, whereas the Masterfile-based projection indicates that 61% of active physicians will be younger than 55 years and 18% will be older than 65 years.

These projections were compared with the Masterfile-based projections of physician supply constructed by HRSA (which used somewhat different meth-

---

©2009 American Medical Association. All rights reserved.
PHYSICIAN WORKFORCE ESTIMATES AND SUPPLY PROJECTIONS

The accuracy of physician supply projections has been questioned because of uncertainty about physician retirement patterns, entry into the profession by US and international medical graduates, and the effect of an increasing number of female physicians.6,10 This is of particular concern given the large cohorts of baby boom physicians who are nearing retirement age. Relative to the Masterfile data, CPS-based projections indicate more young physicians in the workforce and fewer older physicians remaining active. Although these differences had little net effect on projected growth in physician supply, they resulted in projections of a much smaller and younger physician workforce now and in the future.

The differences in workforce estimates between the data sources are likely due to reporting lags in the Masterfile data. Although the Masterfile data gather comprehensive information on new physicians, lags may occur in identifying younger physicians, particularly with increasing numbers of international medical graduates entering the medical profession.4,8,10 A study of Canadian immigrants found that the Masterfile data underestimated migration to the United States with a lag time of 5 years or more.12 Given the large number of young physicians who have been practicing for less than 1 year, even a short reporting lag could account for the 12% discrepancy that our study observed at younger ages in recent years.

Prior work has also documented overcounting of older physicians and attributed this to lags in reporting transitions to retirement.10,11,20 The 3-year frequency at which information is updated for older physicians in the Masterfile data, combined with high retirement rates after age 65 years, could account for the finding that about half as many physicians older than 65 years report being active in the CPS data compared with the Masterfile data.

These findings suggest a number of ways workforce projections based on data from the Masterfile could be made more accurate. Although surveys such as the CPS cannot replace the Masterfile because they lack detail on geography and specialty that is required for workforce planning, they provide benchmark comparisons that can be used to adjust estimates based on Masterfile data. Weights could be constructed for the Masterfile data to ensure that the number of active physicians by age and sex were comparable with estimates from an alternative source such as the CPS. Recent projections by the Association of American Medical Colleges used a survey of older physicians to adjust the 2006 Masterfile data for misreporting of activity status at older ages.5 Alternatively, retrospective information on when each physician changed status (ie, year of licensure for young physicians, year of retirement for older physicians) would allow workforce analysts to construct revised workforce estimates at a later date similar to how the federal government revises economic statistics (such as employment growth) as additional data are reported. Although the most recent year of data would reflect reporting lags, revised data from prior years would not.

Workforce projections are driven by estimates from the past and assumptions about the future. Much of the debate over supply projections has focused on assumptions about the future, such as whether baby boomers will retire earlier than prior generations (or perhaps later because of the current economic downturn) or whether changing lifestyle choices of new physicians will reduce the hours they are willing to work. These are considerations that were not incorporated in our study. However, although uncertainty about the future is important to the workforce debate, this analysis highlights that uncertainty about estimates from the past is also important.

Although this analysis was restricted to physician supply, projections of physician requirements also rely on estimates of the current number of physicians as a starting point for projections. Thus, without more accurate estimates of the size and age distribution of the current workforce, projections of physician supply, requirements, and potential shortages may mislead policymakers as they try to anticipate and prepare for the health care needs of the population.

Author Contributions: Dr Staiger had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis. Study concept and design, analysis and interpretation of the data; or in the collection, analysis, and interpretation of the data; or in the preparation, review, or approval of the manuscript.

Role of the Sponsor: The National Institute on Aging had no role in the design and conduct of the study; in the collection, analysis, and interpretation of the data; in or the preparation, review, or approval of the manuscript.

Financial Disclosures: None reported.

Funding/Support: This work was supported by a grant 2 P01 AG19783-06 from the National Institute on Aging (Dr Staiger).

Statistical analysis: Staiger, Auerbach. Financial Disclosures: None reported.
REFERENCES


