Quality of the Family Physician Component of the AMA Masterfile

P. Tennyson Williams, MD, Michael Whitcomb, MD, and Joseph Kessler

Background: **This investigation was undertaken to gain insight into the validity of the American Medical Association (AMA) Masterfile data.**

Methods: **Allopathic family physicians were chosen as the study population. Omissions were picked up from by comparing the AMA list with the 1990 Ohio Academy of Family Physicians Foundation–Ohio Department of Health (OAFPF–ODH) census. Verification of the 1990 specialty and geographic location of allopathic family physicians not common to both files was achieved by sequentially (1) reviewing the AMA names against 1990 deletions from the 1985 OAFPF-ODH census, (2) contacting physicians directly by telephone, (3) verifying 1990 physician status with county medical personnel, and (4) mailing a brief questionnaire to each physician whose 1990 status remained unverified.**

Results: **The status of specialty and geographic location in 1990 was verified in 91 percent of names not common to both lists. Incorrect omissions (undercounts) and incorrect inclusions (overcounts) offset each other for both lists. Two groups of family physicians contribute to counting biases: family physicians who fulfill short-term goals by part-time practice in several locations, and family physicians who restrict their practice to a limited medical content area.**

Conclusions: **Because of nearly equal offsetting of overcounting (incorrect inclusions) and undercounting (incorrect omissions), the 1990 Ohio family physician AMA Masterfile data is adequate for workforce projections and policy studies when the county data are aggregated at the state level. The overcounting and undercounting for smaller areas or categories must still be studied, however. Application of the AMA Masterfile data of other geographic areas requires a knowledge of the components of undercounts and overcounts of the population being studied. (J Am Board Fam Pract 1996;9:94–9.)**

In the past few years there has been interest in documenting the level of generalist physician supply.[1] Medical workforce studies often lack comparability with regard to inclusion or exclusion of physicians in training, physicians in teaching and research positions, federal physicians, and physicians in administrative positions.[2] In addition, the accuracy of available physician workforce databases is affected by overcounting and undercounting physicians as a result of incorrect specialty designation or practice location.[2-4] Because the AMA Masterfile is the most commonly used source of physician workforce data, we undertook a limited study designed to gain insight into the relative accuracy of the Masterfile data. We present the results of this study.

Submitted, revised, 7 November 1995.

From the Department of Family Medicine, The Ohio State University College of Medicine, Columbus (PTW); the Association of American Medical Colleges, Washington, DC (MW); and the Ohio State University College of Medicine, Columbus (JK). Address reprint requests to P. Tennyson Williams, MD, Department of Family Medicine, The Ohio State University College of Medicine, Room B0902B UCH, 456 W. 10th Avenue, Columbus, OH 43210.

This study was funded by the Fahey Fund of The Ohio State University Development Department. The OAFPF-ODH Census was partially funded by the Ohio Department of Health and the Ohio Academy of Family Physicians.

Methods

Data Sources

AMA Masterfile

The names of nonfederal family physicians and general practitioners in office practice in Ohio were retrieved by county from an early 1991 version of the AMA Masterfile. The AMA Masterfile[5] contains historical data on more than 700,000 active and inactive physicians in the United States. Physicians who attend medical school in the United States are entered into the Masterfile at the time of matriculation. Physicians who do not attend medical school in the United States are entered into the Masterfile at the time they enter US graduate medical education training. Osteopathic physicians are listed if they attend approved allopathic residency train-

ing programs. The Masterfile is updated on a continuous basis by periodic surveys and by information from secondary sources. The file used for this study was a commercially acquired version.

OAFPF-ODH Primary Physician Census

The Ohio Academy of Family Physicians Foundation–Ohio Department of Health (OAFPF-ODH) primary physician census was created in 1975 as a cross-sectional census of generalist physicians in office practice in Ohio and has been compiled every 5 years thereafter. The census is performed by local physician census takers who update a provided list, as described elsewhere.[6,7] A variety of sources (including county medical societies, hospital medical staffs, and telephone directories) are used to select physicians to be deleted (moved, retired, died) from the previous census and to capture names of physicians entering practice since the previous census. Specialties included are allopathic and osteopathic physicians in family and general practice, general internal medicine, and general pediatrics. The 1990 census was conducted from 1 September 1990 through early 1991, after most 1990 residency graduates were in their new practice location and before 1991 residents had graduated. Subspecialists in internal medicine and pediatrics, full-time academic generalist physicians, and physicians in generalist training programs were excluded from this census, consistent with later recommendations of Kindig.[2] For the study reported here, the names of the allopathic family physicians-general practitioners listed in the 1990 census were compared with the names of the allopathic family physicians listed on the AMA Masterfile.

Validation Procedure

All names common to both databases were accepted as accurate without further analysis. All names not common to both the AMA Masterfile and OAFPF-ODH census were analyzed as follows. Osteopathic, resident, and full-time academic family physicians were deleted to assure comparability of the lists. Osteopathic physicians were found by reference to the Ohio Osteopathic Association directory. Physicians who graduated between 1987 and 1990 were deleted on the assumption that they were still in training programs. Full-time academic physicians were determined from American Academy of Family Physicians medical school and residency files and by communication with Ohio academic departments of family medicine offices.

Four sequential steps were taken to determine the 1990 practice specialty status and geographic location of each physician remaining on the list: (1) the original returns of the 1990 OAFPF-ODH census were examined to pick out physicians listed in the AMA Masterfile who had been determined by OAFPF-ODH census takers to have been retired or deceased or to have moved between the 1985 and 1990 enumerations; (2) physicians for whom telephone numbers could be found in directories of the public library were surveyed using a structured telephone interview; (3) local physicians and representatives of county medical societies, academies of family physicians, and local hospitals were queried about listed physicians by telephone and by mail; and (4) the remaining physicians for whom addresses could be found in medical organization directories were mailed a brief questionnaire to indicate their specialty and location status in 1990. Those found to have been incorrectly listed were categorized as being retired, coded in incorrect specialty, otherwise miscoded, moved from the county, or deceased. Finally, because both original lists were segregated by county, duplicate names were sought by alphabetically sorting the valid names without segregation by list source or county. Those found on both lists were removed from the study population and returned to the list of names common to both files. Those for whom no information was determined by any of the above actions were coded "status undetermined." They were not included in the analysis.

Analysis

For each list physician names determined to have been incorrectly listed on either data set were designated incorrect inclusions. Physician names determined to have been listed correctly on the AMA Masterfile but not listed on the OAFPF-ODH census were designated as valid entries for the AMA Masterfile but as incorrect omissions (invalid) from the OAFPF-ODH census. Those listed correctly on the OAFPF-ODH census but not listed on the AMA Masterfile were designated as valid entries for the OAFPF-ODH census but as incorrect omissions (invalid) from the AMA Masterfile. The following formulae were used:

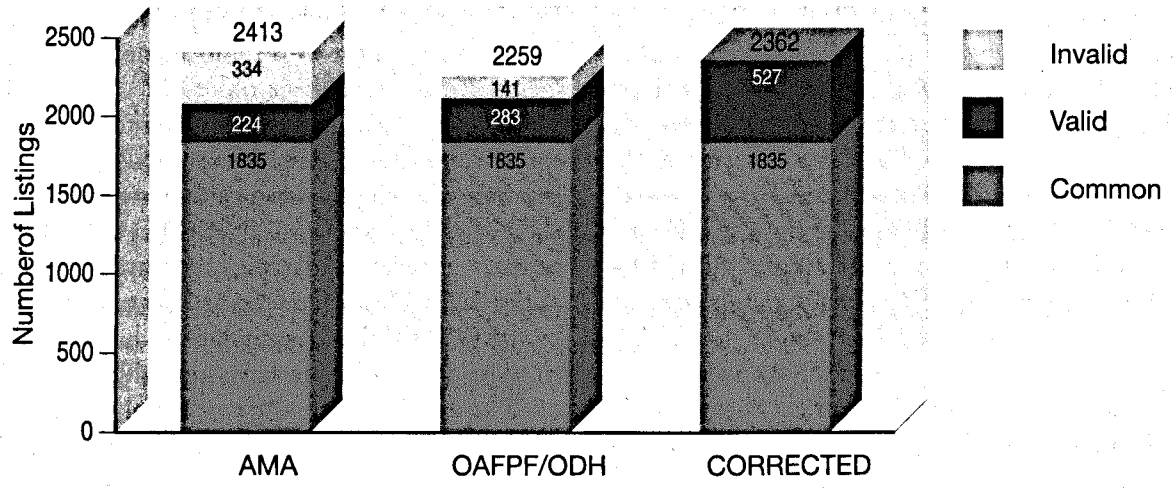

Figure 1. Valid and invalid listings for the AMA Masterfile, the OAFPF-ODH census, a final corrected census. AMA - American Medical Association; OAFPF-ODH - Ohio Academy of Family Physicians Foundation–Ohio Department of Health.

Valid listings = Names common to both files + correct listings

Invalid listings = Incorrect omissions + incorrect inclusions

Percent total invalid = (Incorrect inclusions + incorrect omissions)/(final correct count) × 100

Percent net invalid = (Incorrect omissions − incorrect inclusions)/(final correct count) × 100

A positive net invalid number constitutes undercounting, and a negative net invalid number constitutes overcounting.

Results

The original AMA Masterfile and OAFPF-ODH census contained a total of 3013 names after osteopathic physicians, resident physicians, and full-time academic physicians were deleted from the two lists: 671 from the AMA Masterfile and 6 from the OAFPF-ODH census. Of the 3013 names, 1835 were common to both lists. The 1178 allopathic family physician names not common to both lists were those for whom validation was sought for their 1990 geographic location and specialty. We were unable to validate 9 percent of the names not common to both files. The specialty and geographic location of about 80 percent of physicians not common to both files were validated from the original reports of the OAFPF-ODH census takers and directly from the individual physicians by telephone calls or mailed questionnaires. The 11 percent remaining validations were obtained from secondary sources in the counties of record. Figure 1 illustrates the relative size of the valid, invalid, and common components found for the AMA Masterfile, the OAFPF-ODH census, and the final corrected census (2362).

Fifty-four percent of invalid entries of the AMA Masterfile and 37 percent of invalid entries of the OAFPF-ODH census were due to incorrect inclusion (Table 1). The net invalid number of the AMA Masterfile was −51 physicians (incorrect inclusions exceeded incorrect omissions), an overcount of 2.2 percent of the final corrected census. The net invalid number of the OAFPF-ODH census was +103 (incorrect omissions exceed incorrect inclusions), an undercount of 4.4 percent of the final corrected census.

The three most frequent reasons found for incorrect inclusion were retirement, specialty errors, and miscoding. These factors accounted for 89 percent of the total incorrect inclusions (87 percent of AMA Masterfile and 93 percent of the OAFPF-ODH census, Table 2). Specialty errors resulted largely from physicians who were family physicians by training or certification, but who were limiting their practice to components of family medicine, such as school health, nursing home practice, emergency medicine, or occupational medicine. These represented both recent career changes and physician preference for categorizing (coding) their restricted practice. The miscoded designation included undetermined

Table 1. Entry Characteristics of Valid and Invalid Findings.

Characteristics	AMA No. (%)	OAFPF-ODH No. (%)
Valid, common to both files	1835	1835
Valid, on this file, not on other	244	283
Invalid, incorrect omissions	283 (46)*	244 (63)*
Invalid, incorrect inclusions	334 (54)†	141 (37)†
Total invalid entries‡	617	385
Invalid entries (%)§	26.1	16.3‖
‡Net invalid entries	(-)51	103

AMA - American Medical Association; OAFPF-ODH - Ohio Academy of Family Physicians Foundation - Ohio Department of Health.
*Percent of invalid entries due to incorrect omissions.
†Percent of invalid entries due to incorrect inclusions.
‡Names common to both lists were assumed to be valid. Therefore total errors is the sum of invalid incorrect omissions and invalid incorrect inclusions.
§Percent invalid entries = total invalid entries of list divided by 1835 + 244 (valid on AMA list) + 283 (valid on OAFPF-ODH list) [2362] x 100.
‖ $\chi^2 = 53.25$; df = 1; $P < 0.00001$ (McNemar chi-square).

reasons, error in designation as allopathic or osteopathic physician, and out-of-state location.

Because reasons for incorrect omission were beyond the knowledge of individuals or county secondary sources, they could not be described.

When the names of each file were aggregated and re-sorted alphabetically, a small number of physicians' names were found to be common to the AMA Masterfile and the OAFPF-ODH census. In every case one file had them listed in a county different from the other file. These duplications usually represented a move of the physicians's office from one county to another within the target year of the study, though some duplications reflected offices in each of two adjacent counties. These names were returned to the "common to both files" category before analysis. A sensitivity analysis listing these duplications as incorrect inclusions due to incorrect location instead of inclusion in the common group did not significantly alter the statistical findings.

Discussion

Undercounting and overcounting errors existed for both the AMA Masterfile and the OAFPF-ODH census, and were nearly offsetting for each file. The AMA process of acquiring data from medical schools and residency programs assures a high degree of accuracy in initial entry into the Masterfile, whereas the dependence upon secondary sources of information for updating tends to delay deletions as a result of retirement, moving, or death. This delay in deletion increases incorrect inclusions of physicians who move, retire, or die, leading to overcounting. In a much earlier study, Cherkin and Lawrence[3] reported a delay of approximately 2 months for the removal of names of deceased physicians. In our experience it was not uncommon to encounter longer interval delays, occasionally as long as 10 years.

The finding of 2.2 percent net error in the AMA Masterfile is consistent with two previous comparisons: a 1974 on-site inventory of physicians in two sample counties in Massachusetts (2.0 to 2.3 percent deviation),[8] and a survey of practicing physicians in the state of Washington.[3] In the latter study 99.9 percent of physicians licensed and living in Washington State were known to the AMA Masterfile, and 94 percent were living in Washington State. Most of the discrepancies of this study were the result of the time lag between a location change and subsequent updating of the Masterfile.

A major source of overcounting in this study was specialty error. Specialty error is a function of self-designation. The lack of definitions for subspecialty areas in which family physicians often choose to limit their practice leaves little choice but to designate the specialty of their training. Specialty errors accounted for 149 instances of incorrect inclusion; about one third of all reasons for incorrect inclusion for both lists, and 2.0 percent (OAFPF-ODH) to 4.3 percent (AMA) of the final correct census. Because family physicians who sub-

Table 2. Reasons for Incorrect Inclusions, Final Census.

Reasons	AMA No. (%)	OAFPF-ODH No.(%)
Retired	133 (5.6)	57 (2.4)
Specialty error	101 (4.3)	48 (2.0)
Miscoded*	58 (2.5)	26 (1.1)
Moved from county	38 (1.6)	5 (0.2)
Deceased	4 (0.2)	5 (0.2)
Total	334 (14.2)	141 (5.9)

AMA - American Medical Association; OAFPF-ODH–Ohio Academy of Family Physicians Foundation - Ohio Department of Health.
*Miscoded as to medical doctor or doctor of osteopathy status, reason undetermined, out-of-state location.

specialize might no longer fulfill generalist roles in the health care system, they constitute a source of overcounting bias of as much as 4 percent of the family physician component of the generalist workforce. If the trend continues, as it seems to be, the balance between overcounting and undercounting might not continue. This component is an independent source of error, because it is not subject to delay of entry error. Future workforce studies must address how these physicians are to be detected and categorized if the current level of accuracy is to continue or improve.

Other sources of generalist workforce underestimation or overestimation were not addressed in this study, but they include physicians who practice part-time in one location, the midlevel practitioners who perform generalist functions, and physicians listed as specialists but who practice primarily as generalists. The problem of determining full-time equivalents is greatest when assessing the generalist physician workforce, as it includes both the overcounting of part-time physicians and the undercounting of the generalist component of specialists and allied health professionals.

We did not determine the reasons for errors of omission (incorrect omission) because that information was not known to our sources of information. Entry into the AMA Masterfile has been found to be quite accurate; therefore, the finding of omissions was unexpected. These omissions must occur because of listings in other locations (such as undetected moves) or in other specialties (inappropriate self-designations or entry coding error), or incorrect deletions. Incorrect deletions seem unlikely, as they would occur only when incorrect information is gathered from secondary sources, though erroneous assumption of retirement, such as after resignation from hospital staff or medical organization, might occasionally occur. Cherkin and Lawrence,[3] in comparing the AMA Masterfile with Washington State licensing data, found 274 of the 302 physicians listed on the licensing file but missing from the AMA Masterfile mostly located in other states. We did not develop a method of capturing or categorizing these omissions. They were discovered by sorting for physician names on the OAFPF-ODH census that were not listed on the AMA Masterfile.

Physicians working full time but in several part-time positions in different geographic locations constitute a source of error that might cause allocation of either location or specialty or both incorrectly. Many of these physicians work at several part-time salaried positions to fulfill short-term goals. They could either be counted in several counties (overcounted) or be missed entirely (undercounted). Working in both specialist settings (emergency facilities) and generalist settings (community clinics, urgent care centers) would create errors in specialty designations. Most of these physicians are not members of a local county medical society or hospital medical staff and therefore are not socialized into the local medical community. Consequently, they are often overlooked by local census takers and by secondary sources upon which the AMA Masterfile depends for status change information. This component of the physician workforce might be increasing as residency graduates use such opportunities to test several career choices before making final career decisions.

There are several limitations to this study. First, the 1991 AMA Masterfile data were taken from a commercial copy originally obtained in early 1991, whereas the 1990 OAFPF-ODH study was begun in September 1990 and completed early in 1991. Both files, however, were derived after the 1990 class of resident graduates had selected practice sites and before the 1991 class had done so. Second, we did not check the validity of the 1835 names common to both lists. Nevertheless, potential errors in this component of the lists would be common to both lists and would be expected to reflect the same offsetting of incorrect omissions and incorrect inclusions found in the study components. Consequently, there should be minimal effect on overall accuracy and net results. Third, we excluded from the analysis all physicians whose status was unaccounted for by any of the study strategies. These physicians constituted 9 percent of the "not common" group, and 5.8 percent of the entire study population. Physicians whose specialty and practice location remained unknown were unlikely to have been in practice in Ohio at the time of this resurvey (spring of 1994). Though their status in 1990 was unknown, they would be overcounted more often than correctly listed. Finally, it is highly probable that there were physicians missing from both lists.

Comparison of other state databases with the AMA Masterfile is necessary to confirm the find-

ings of this study. Nevertheless, these findings suggest that both 1990 Ohio databases are sufficiently accurate to direct medical workforce policy formation when the county data are aggregated to the state level. Additional study is necessary to determine whether the same accuracy holds true for segregated data, such as rural-urban differences, and for other specialties. Improvement of the accuracy of family physician workforce databases would occur by correct allocation of the work of itinerant family physicians and of family physicians who have limited the scope of their practice.

This study confirms the usefulness of the AMA Masterfile, providing that unwanted categories of physicians are first carefully deleted. An important finding is the two components of the family physician workforce that introduce bias into current databases: physicians who work part-time at several locations doing different categories of work, and physicians who limit their practice to areas that no longer reflect the roles of family physicians. Further study of these components should be conducted to determine their impact on the generalist physician workforce.

References

1. Weiner JP. Forecasting the effects of health reform on US physician workforce requirement. Evidence from HMO staffing patterns. JAMA 1994;272:222-30.
2. Kindig DA. Counting generalist physicians. JAMA 1994;271:1505-7.
3. Cherkin D, Lawrence D. An evaluation of the American Medical Association's Physician Masterfile as a data source—one state's experience. Med Care 1977; 15:767-9.
4. Weiner JP, William MK, Landis WB. An analysis of the current and projected need for primary care physicians in Maryland. Md State Med J 1984;33:732-5, 740-6.
5. Roback G, Randolph L, Seidman B. Physician characteristics and distribution in the U.S. 1993 edition. Chicago: American Medical Association, 1993.
6. Brown LM, Williams T. Survey of Ohio primary physicians: a projection of future needs to the year 2000. Ohio State Med J 1978;74:29-33.
7. Brown LM, Williams T. A resurvey of Ohio family physicians. A projection of future needs to the year 2000. Ohio State Med J 1982;78:117-20, 23.
8. Family physician manpower survey. Boston: Public Policy Committee, Massachusetts Academy of Family Physicians, 1974.