Radiology Groups’ Workload in Relative Value Units and Factors Affecting It1

**PURPOSE:** To measure diagnostic radiology groups’ workload in physician work relative value units (RVUs) and identify factors affecting it.

**MATERIALS AND METHODS:** In 1996 and 1997, the authors surveyed diagnostic radiology and radiation oncology groups regarding finances, workload, and basic characteristics. The study was based on approximately 100 diagnostic radiology groups. The authors analyzed the distribution of workload in different categories of groups, conducting multiple statistical analyses.

**RESULTS:** The annual numbers of procedures were approximately 10%–15% lower than those in a comparison survey with a good response rate. The annual number of RVUs per full-time equivalent (FTE) diagnostic radiologist was highly variable in every group category, as was the number of RVUs per clinical work hour. Multivariate regression analysis indicated that variation in the annual number of hours worked did not explain variation in annual workload. RVUs per FTE radiologist were higher the greater the percentage of a group’s workload from interventional, computed tomographic, and magnetic resonance imaging procedures.

**CONCLUSION:** Given the likely response bias, the annual workload per FTE radiologist probably averaged approximately 4,000 RVUs in academic groups and approximately 6,000 in nonacademic groups, but the large, unexplained variance means the average values should not be taken as norms.

“What is the workload of a typical radiologist?” is the question most frequently asked of the American College of Radiology (ACR) research department. Radiology practices want to know the answer to plan their staffing and operations. In addition, they seek the answer to use in response to outside entities such as hospital administrators that pressure them to increase or decrease their staffs.

Extensive information on radiologists’ workload measured by using the number of procedures performed annually (1–4) is available. However, the value of this information is limited by the reality that a procedure is not a uniform unit of work. At the extreme, an interventional procedure can involve as many hours of a radiologist’s time as the simplest conventional film hard-copy study requires minutes (5).

A relative value unit (RVU) is a metric designed to measure physicians’ work activity in uniform units. Since the development of the Medicare radiology relative value scale (RVS) under the auspices of the ACR in the late 1980s (6), a carefully and thoughtfully constructed RVS for measuring radiology workload has been available. However, to our knowledge, only limited information on workload measured in RVUs has been available (1,4,7,8), principally because few practices have been able to furnish the requisite detailed data.

In this article, we seek to remedy this information deficit by providing much more extensive information on workload measured in RVUs than has heretofore been available. We first present data on the annual number of RVUs per full-time equivalent (FTE) radiologist for radiology groups generally and for different group types (ie, academic versus nonacademic), group sizes, and geographic areas. We present not only average and median values but also the values in the 25th and 75th percentiles to show how much—or little—variation there is. Because relatively few groups could supply the requisite data, we
also tried to estimate how typical our data are of the general range of radiology groups in the United States. Next, we investigated whether, as anticipated, annual workload measured in RVUs per FTE radiologist is less variable than annual workload measured in procedures per FTE radiologist. In this investigation, we used both the coefficient of variation (CV), which is a standard statistical measure of variability, and multivariate regression analysis. During our research, advisors suggested that the annual number of RVUs is likely to be substantially affected by (a) the fact that groups vary substantially in their annual number of work hours and (b) a group’s mix of procedures, with some types of radiologic work generating considerably more RVUs per hour worked than others. We therefore investigated whether taking either of these factors into account makes measured workload less variable. For this task we again studied CVs and performed multiple regression analysis.

Finally, we used multivariate analysis—specifically, multiple regression analysis—to ascertain what factors, such as group type, group size, and hours worked annually, independently affect annual workload when the effects of all the other factors are considered simultaneously. We present a formula based on this analysis with which a group can calculate the average workload of groups that have the same characteristics that it has.

MATERIALS AND METHODS

Data Sources

In 1996, we surveyed a stratified random sample of 320 diagnostic radiology and radiation oncology groups in the United States by mail, asking them for substantial financial information (almost entirely about outlays rather than revenues), about the basic characteristics of the group (such as the number of radiologists), and about the details of their workload. Stratification was by group size and region of the country. For workload, we asked the groups to supply the number of procedures performed in the most recent year for which data were available and to provide the count separately for each physicians’ CPT (Current Procedural Terminology, copyright annually by the American Medical Association) code, and within each CPT code, separately for global (nonhospital) and professional component (hospital) procedures. Despite up to three repeated mailings at intervals of 1–2 months, only 73 (23%) groups responded with the requested workload data. To increase the number of groups for which we had data, in mid-1997, we mailed the survey to 490 diagnostic radiologists and radiation oncologists who had leadership positions in the ACR or its state chapters, asking them to provide the requested data for their group. We received 77 (16%) additional replies regarding workload. The number of replies in 1997 was depressed by the absence of follow-up mailings, short time lines, instances in which multiple addressees were members of the same group, and instances in which the addressees were members of a group that had already responded in 1996. Of the total of 150 responding groups that supplied workload data in 1996 or 1997, two were deleted because of procedure coding inconsistencies, and 37 were radiation oncology only groups or combined oncology and diagnostic groups and thus had to be omitted from our analysis. (The ACR listing of groups does not distinguish those that include radiation oncologists, so they were included in the survey.) Thus, we could have potentially analyzed data from 111 groups. Four of these did not supply information on the group type (ie, academic, nonacademic private radiology, nonacademic multispecialty, or nonacademic government) and thus are omitted from all tables. In addition, data on the three two-member groups that responded are omitted from the tables containing weighted data because these groups contained too few members for the weighting process (described later in Analysis subsection). In late 1998, we attempted to contact by telephone or fax all 111 diagnostic-only groups for which we had data on workload in RVUs to learn about their work hours. We asked about the number of clinical hours worked by a typical full-time group member in a full week and the average number of days taken annually for vacation and for continuing medical education and professional meetings by a full-time group member. We obtained information for 88 (79%) of the 111 groups. These 88 groups minus the three two-member groups constituted the 85 groups included in our analysis of variability. One of the 88 groups did not provide information on its type; the remaining 87 are the subject of the regression analyses.

As previously reported (1), in its 1996 Survey of Hiring by Groups, the ACR contacted a stratified random sample of 794 groups by mail, asking them (among other things) about workload measured in procedures per year. This survey had a 78% response rate (617 valid responses and seven groups indicating they no longer existed). We compared information from the 1996 hiring survey, in which the nonresponse bias presumably was low because of the high response rate, with data from our survey measuring workload in RVUs to illuminate the response bias in the latter.

Measurements and Definitions

RVUs were measured in physician work RVUs of the 1998 Medicare resource-based RVUs (RBRVS) (9). Readers who are not familiar with the RBRVS may not be acquainted with physician work RVUs and may be more familiar with the concept of professional component RVUs. With the RBRVS of 1998, physician work RVUs for a radiologic procedure, the measure we used, equaled two-thirds of the total RVUs assigned to the professional component of the procedure. Physician work RVUs generally will not change in the RBRVS in the years subsequent to 1998, but the total professional component RVUs will.

We measured the FTE number of radiologists in a group as the group’s number of full-time members plus the appropriate proportional count of each part-time member of the group. For example, a half-time radiologist counted as 0.5 FTE.

The number of clinical hours worked annually per FTE radiologist was estimated by (a) subtracting the reported days off for vacation and continuing medical education or professional meetings from the total number of workdays in a year (assuming a 5-day work week); (b) subtracting an additional 12 days off for holidays, illness, and other absences (10); (c) dividing by five to obtain the weeks worked, and (d) multiplying by the reported number of clinical work hours in a full week. On the survey question about clinical hours worked weekly, the respondents were instructed to exclude administrative, teaching, and research time unless these involved usual clinical work.

Advisors to our study indicated that computed tomographic (CT), magnetic resonance (MR) imaging, and interventional or angiographic procedures generate more RVUs per hour than do other procedures, and that this should substantially affect the amount of work radiologists accomplish when work is measured in RVUs. (The RBRVS deliberately is not simply a measure of time, but rather it also includes “intensity”—defined as RVUs per unit of time—to reflect the fact...
that the work involved in procedures also varies in required mental effort, stress, physical effort, training, etc. Because of what the advisors told us, we called CT, MR imaging, and interventional or angiographic procedures “high productivity” procedures and measured their role in a group’s workload.

Analysis

In most analyses, data from each group were weighted to make the tables representative of all diagnostic radiologists in the designated categories in the United States. The weighting method used has been explained in detail elsewhere (2). Briefly, we stratified our sample population into four size categories—groups of three to four radiologists, five to seven radiologists, eight to 10 radiologists, and 11 or more radiologists—from each of the four census regions (Northeast, Midwest, South, and West), for a total of 16 sampling strata. To make the responding groups representative of all groups in the United States, the responses from each stratum were weighted by using the estimated number of groups in that stratum in the United States divided by the number of responses in the stratum. To produce a final weight that made the data representative of all diagnostic radiologists (rather than groups) in the country, the initial weight described in the previous sentence was then multiplied by the number of radiologists in the group (in FTEs).

In exploring the variability of different measures of workload, we used the CV as the main measure of variability. The CV, which is defined as the SD divided by the mean (average), is the most common statistical measure of variability.

Multivariate analysis was conducted by using ordinary least squares regression. In the principal multivariate analyses, data were not weighted because the number of responding groups was inadequate. In regression analysis, it is a rule of thumb that there should be a minimum of approximately 20 cases per independent variable. Nonweighted, our data set was just adequate from this standpoint. Weighting has a statistical effect that is equivalent to reducing the number of cases in the data set. Thus, in the principal multivariate analyses, each group was simply counted as one. However, the same regression analyses were performed in a weighted manner; a summary of the results is reported in the text but not shown in the tables.

All data analyses were carried out with SAS (Statistical Analysis System; SAS Institute, Cary, NC) software. “Significant,” as used in the text, was defined as a P value of less than or equal to .05 by a two-tailed test. In the cases in which we made comparisons among more than two categories in a classification—for example, in the comparison of the four practice size categories—the threshold P value was adjusted with the Bonferroni inequality correction (11) to ensure that the probability that any one or more of the possible pairwise comparisons was significant was less than or equal to .05. Specifically, for a comparison of four categories, this means that the probability associated with an individual pairwise comparison must be less than or equal to .008.

RESULTS

Data Validation

One-tenth of the responses to our survey were from groups that contained two to four radiologists, whereas according to results of the ACR 1995 Survey of Radiologists (12), more than 40% of groups that perform diagnostic radiology are this size. These small groups were less likely to respond to the survey than were the other groups; if response had been proportional, then the 1996 portion of the survey alone would have produced more responses from these groups than we received in total. However, the exact magnitude of this response bias is unknown because we do not know the size distribution of the groups contacted in 1997.

A comparison of the last two columns in Table 1 reveals that for all groups that contained five or more radiologists, our survey results indicated an estimated annual number of procedures per FTE radiologist (9,800) that was 12% lower than the corresponding datum (11,200) from the reference survey, which was relatively free of response bias. The difference was statistically significant. Similar differences were found for many of the specific categories of groups in Table 1, although some of these differences—for example, that for academic groups—were not statistically significant because of the smaller sample size.

In contrast, there was good agreement between the current survey’s finding of an average of 0.50 physician work RVUs per procedure and the only comparison statistic that we developed. This statistic is procedures performed on Medicare patients by radiologists in 1995, which averaged 0.75 professional component RVUs (1), which corresponds to 0.50 physician work RVUs. Overall, it seems wise to regard the data from our survey as probably an underestimation of the workload by 10%-15% owing to response bias.

Workload in RVUs per FTE Radiologist per Year

Our survey results indicated that diagnostic radiologists in groups containing five or more radiologists averaged 4,800 physician work RVUs per FTE radiologist annually (Table 1). There was much variability. Twenty-five percent of radiologists in groups with five or more members were in groups in which the annual workload averaged 3,200 RVUs or less, whereas 25% were in groups in which it averaged 6,100 RVUs or more (Table 1, 25th and 75th percentile columns).

Diagnostic radiologists in primarily academic groups containing five or more radiologists averaged 3,500 RVUs annually, which was two-thirds as much as the average annual number of RVUs of radiologists in nonacademic groups with five or more members—5,200. Among nonacademic groups, there were no statistically significant differences in workload based on group size, region of the country, or whether the practice included some nonhospital offices or was a hospital-only group.

Within every category of group, there was much variability in workload. For example, one-fourth of radiologists in the nonacademic groups that contained five to seven radiologists were in groups in which the average annual workload was 3,800 RVUs or less, whereas at the other extreme, one-fourth were in groups in which the workload averaged 6,300 RVUs or more. In half the categories shown in Table 1, as in this example, the workload at the 75th percentile exceeded the workload at the 25th percentile by at least 65%.

In academic groups, the number of RVUs per procedure averaged 0.56, which was significantly more than the average number of RVUs per procedure for the nonacademic groups—0.48. The only other statistically significant difference in average number of RVUs per procedure was between the nonacademic groups of five to seven radiologists (0.44 RVUs) and nonacademic groups of 11 or more radiologists (0.50 RVUs).

Variability of Different Measures of Workload

As explained in the Introduction section, the annual workload per FTE diag-
nastic radiologist was expected to be less variable when it was measured in RVUs than when it was measured in procedures, and still less variability was expected when workload was measured in RVUs per clinical hour worked or an adjustment was made for a group's mix of procedures. The data testing these expectations are presented in Table 2.

The results of analysis of the entire set of 85 groups for which we had the necessary data showed very similar variability, regardless of whether workload was measured in annual number of procedures per FTE radiologist, annual number of RVUs per FTE radiologist, or RVUs per clinical hour worked. Specifically, the CV for all three measures was between 0.39 and 0.41. For the nonacademic groups, the only category for which we had a fairly substantial number of cases \( (n = 71) \) and therefore relatively reliable statistics, the CV was also similar—between 0.36 and 0.39—among the three measures of workload.

When we compared the nonacademic groups in which greater than 41% to less than 48% of their RVUs were from high productivity procedures (on a weighted basis, this is approximately the middle third of the observed range in mix of procedures) with all the nonacademic groups, the former group was less variable in annual number of RVUs per FTE radiologist, as expected. The CV for the former group was 0.25 compared with 0.36 for all the nonacademic groups. Similarly, this middle third was more homogeneous in RVUs per clinical hour worked: The CV was 0.30 compared with 0.39 for all the nonacademic groups.

**Factors Affecting Workload**

The results of multivariate analyses of the factors affecting workload are presented in Tables 3 and 4. Multivariate analyses are used to estimate the independent effect each factor has after controlling for the effects of all the other factors included in the analysis. The left-hand third of Tables 3 and 4 contains the results of analysis in which the annual number of procedures per FTE radiologist was the measure of workload.

The one statistically significant finding was that workload in academic groups was, as expected, lower than that in otherwise similar nonacademic groups—by approximately 2,200 procedures. However, contrary to expectations, the analysis results indicated that the annual number of hours worked had no statistically significant effect on the annual number of procedures, and neither did the proportion of the group’s procedures that were in the high productivity category, which consisted of MR imaging, CT, and interventional procedures (Table 4). Because these procedures are relatively time-consuming, we expected that the groups in which the work mix included a relatively large portion of these procedures would perform fewer procedures annually per radiologist than would the otherwise similar groups. Moreover, both these unexpected findings ruled out the notion that there was an effect of the variable in

<table>
<thead>
<tr>
<th>TABLE 1</th>
<th>Workload in Various Types of Groups</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group Characteristic</td>
<td>Unweighted No. of Groups</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>All groups of five or more radiologists</td>
<td>96</td>
</tr>
<tr>
<td>No. of radiologists in group (nonacademic groups only)</td>
<td></td>
</tr>
<tr>
<td>3–4</td>
<td>8</td>
</tr>
<tr>
<td>5–7</td>
<td>21</td>
</tr>
<tr>
<td>8–10</td>
<td>19</td>
</tr>
<tr>
<td>≥11</td>
<td>38</td>
</tr>
<tr>
<td>Practice type†</td>
<td></td>
</tr>
<tr>
<td>Primarily academic</td>
<td>18</td>
</tr>
<tr>
<td>Nonacademic private, radiology only</td>
<td>74</td>
</tr>
<tr>
<td>Practice location‡</td>
<td></td>
</tr>
<tr>
<td>Hospital only</td>
<td>13</td>
</tr>
<tr>
<td>Both hospital and office</td>
<td>63</td>
</tr>
<tr>
<td>Region§</td>
<td></td>
</tr>
<tr>
<td>Northeast</td>
<td>16</td>
</tr>
<tr>
<td>Midwest</td>
<td>22</td>
</tr>
<tr>
<td>South</td>
<td>28</td>
</tr>
<tr>
<td>West</td>
<td>12</td>
</tr>
</tbody>
</table>

Note.—Groups of two radiologists are omitted in the Table. The numbers in parentheses are the standard error.

* Data are the mean number of RVUs.
† Data are the mean number of procedures.
‡ Statistically significant difference between current survey and reference survey \( (P < .05) \).
§ Groups of five or more radiologists. Does not include nonacademic governmental and multispecialty groups.
¶ Nonacademic groups of five or more radiologists. Does not include office-only practices.

* * *
question, but the effect was not quite statistically significant. The associated probability for each finding was approximately .70, whereas to be counted as statistically significant, we required a *P* value of less than or equal to .05. Not quite statistically significant (*P* = .07) was the finding that workload decreases with group size, other things (including procedure mix and academic or nonacademic status) being equal.

Overall, the four independent variables in Table 4 accounted for only 18% of the variance in the annual number of procedures per radiologist among the groups. (The *R*<sup>2</sup> statistic in Table 3 is defined as this standard error). Nonetheless, the overall regression equation was highly significant (*P* = .002).

In a variant (not shown), we used the proportion of interventional procedures rather than the proportion of high productivity procedures as an independent variable, reasoning that interventional procedures are clearly the most time-consuming. The results were similar to the regression shown; this procedure mix variable was nowhere near statistically significant, and *R*<sup>2</sup> was equal to 0.18.

The middle third of Tables 3 and 4 contains results of a regression analysis in which the annual number of physician work RVUs per FTE radiologist was the measure of workload. As expected, the workload in academic groups was lower than that in otherwise similar nonacademic groups—by approximately 1,800 RVUs. Procedure mix also mattered in the expected manner. Specifically, for each additional percent of a group’s total number of RVUs that originated from high productivity procedures, the annual workload averaged approximately 64 RVUs greater, other factors being equal. The annual number of hours worked did not have a significant effect. Moreover, its effect was nowhere near significant. Group size had an effect that neared statistical significance (*P* = .06); the larger groups had apparently greater workloads per FTE radiologist, other things being equal.

Overall, the four independent variables accounted for 24% of the variance in annual number of RVUs per radiologist; this was only a moderately higher percentage than that calculated when workload was measured in annual number of procedures per radiologist. Thus, again, the estimates of a group’s actual workload that were produced by using these four characteristics were not particularly accurate. Specifically, the SD of the difference between the estimated value and the actual value was approximately 1,640 RVUs annually, a full one-third of the average actual value. (This one-third is the CV in Table 3). Nonetheless, overall, the regression was highly significant (*P* = .001).

We carried out variants of this regression analysis (not shown) by using different measures of procedure mix. In one variant, an additional “medium productivity” independent variable, defined as the proportion of RVUs from ultrasonographic procedures, was added. It was not statistically significant, and *R*<sup>2</sup> remained 24%. In another variant, we used separate independent variables for the proportion of RVUs from CT procedures, from MR imaging procedures, from nuclear medicine procedures, and so forth. In general, these variables were not statistically significant. *R*<sup>2</sup> increased to only 0.27, which is a smaller increase than that which would be expected solely as a result of adding the same number of purely random, independent variables. (The adjusted *R*<sup>2</sup> statistic [not shown] is used to measure the extent to which the *R*<sup>2</sup> of the actual independent variables exceeds the *R*<sup>2</sup> that would be expected from an equal number of purely random, independent variables.)

The results of analysis in which the natural logarithm of a group’s annual total (not per radiologist) number of RVUs was the measure of workload are presented in the right-hand third of Tables 3 and 4. (Analyses in logarithmic form have certain convenient features.) The effect of group size was highly significant: Each 1% increase in group size was associated with a 0.96% increase in total workload, other things being equal. Academic status had a strongly significant effect: The academic groups produced only 62% as many

---

**TABLE 2**  
Variability of Different Measures of Workload

<table>
<thead>
<tr>
<th>Workload Measure and Category of Groups</th>
<th>Unweighted No. of Groups</th>
<th>Mean</th>
<th>SD</th>
<th>CV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Annual procedures per FTE radiologist (in thousands)</td>
<td>All groups</td>
<td>85</td>
<td>9.25</td>
<td>3.75</td>
</tr>
<tr>
<td>Ac</td>
<td>Academic groups</td>
<td>14</td>
<td>6.34</td>
<td>2.45</td>
</tr>
<tr>
<td>N</td>
<td>Nonacademic groups</td>
<td>71</td>
<td>9.99</td>
<td>3.67</td>
</tr>
<tr>
<td>Annual physician work RVUs per FTE radiologist (in thousands)</td>
<td>All groups</td>
<td>85</td>
<td>4.83</td>
<td>1.88</td>
</tr>
<tr>
<td>A</td>
<td>Academic groups</td>
<td>14</td>
<td>3.51</td>
<td>1.25</td>
</tr>
<tr>
<td>N</td>
<td>Nonacademic groups</td>
<td>71</td>
<td>5.16</td>
<td>1.86</td>
</tr>
<tr>
<td>N</td>
<td>Nonacademic groups with &gt;41% to &lt;48% of RVUs from “high productivity” procedures</td>
<td>25</td>
<td>5.98</td>
<td>1.52</td>
</tr>
<tr>
<td>A</td>
<td>Physician work RVUs per clinical work hour</td>
<td>All groups</td>
<td>85</td>
<td>2.67</td>
</tr>
<tr>
<td>A</td>
<td>Academic groups</td>
<td>14</td>
<td>2.32</td>
<td>0.78</td>
</tr>
<tr>
<td>N</td>
<td>Nonacademic groups</td>
<td>71</td>
<td>2.76</td>
<td>1.08</td>
</tr>
<tr>
<td>N</td>
<td>Nonacademic groups with &gt;41% to &lt;48% of RVUs from “high productivity” procedures</td>
<td>25</td>
<td>3.10</td>
<td>0.93</td>
</tr>
</tbody>
</table>

---

**TABLE 3**  
Results from Regression Analyses: Summary Statistics

<table>
<thead>
<tr>
<th>Summary Statistic</th>
<th>Annual Procedures per FTE Radiologist</th>
<th>Annual Physician Work RVUs per FTE Radiologist</th>
<th>Natural Logarithm of Group’s Annual Total RVUs*</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>R</em>&lt;sup&gt;2&lt;/sup&gt;</td>
<td>0.180</td>
<td>0.237</td>
<td>0.715</td>
</tr>
<tr>
<td>Root mean squared error</td>
<td>3,245</td>
<td>1,639</td>
<td>0.447</td>
</tr>
<tr>
<td>CV</td>
<td>0.34</td>
<td>0.33</td>
<td>0.04</td>
</tr>
<tr>
<td>P of entire equation</td>
<td>.002</td>
<td>.001</td>
<td>.001</td>
</tr>
</tbody>
</table>

* In the logged regression analysis, all independent variables are also the natural logarithms, with the exception of academic status.
RVUs as did the otherwise similar nonacademic groups. (0.62 is the antilog of −0.477, the coefficient of the academic status variable.) The proportion of RVUs from high productivity procedures also was significant: Each 1% increase in this proportion (for example, an increase from 0.45 to 0.4545) was associated with a 0.35% increase in workload, other things being equal. The number of hours worked annually had no statistically significant effect.

Overall, the four independent variables accounted for 72% of the variance in the groups’ annual total number of RVUs; this was a much higher percentage than that in the other regression analyses. This finding plus the very strong statistical significance of group size indicated, unsurprisingly, that group size was the dominant factor in the total number of RVUs a group performed annually. Nonetheless, this regression did not provide a more accurate estimate of a group’s RVUs than did the regression that estimated the number of RVUs per radiologist (central third of Tables 3 and 4). Calculations of the root mean squared error statistic of the regression for total number of RVUs (this was necessary to undo the logarithmic form) indicated that the total number of RVUs predicted for a group by using the regression equation typically differed from the actual value by one-third to one-half.

The results of regressions weighted (not shown), as described earlier (Materials and Methods section), to represent all diagnostic radiologists in the United States were very similar to those of the three regressions in Tables 3 and 4. Most important, the typical error in estimating a group’s workload from the regression (i.e., the root mean squared error) was very similar, as was the CV. In addition, the effect of academic status was uniformly significant and broadly similar in magnitude to that shown in Table 4, and the effect of hours worked annually never approached statistical significance. However, unlike the results in Table 4, the effect of procedure mix was usually only marginally statistically significant, whereas that of group size generally was highly significant.

### DISCUSSION

#### Other Information Sources

To our knowledge, the previously available sources on diagnostic radiologists’ workload measured in RVUs were few and covered limited types of groups. However, these sources had good response rates. Conoley and Vernon (4) reported data on 1989–1990 workloads in 19 large multispecialty clinics. (Here, “clinics” refers to practices like Cleveland Clinic and Mayo Clinic.) They used an earlier RVS and different measure of FTE radiologists, but when their values are converted to our metrics, their average value becomes 4,440 RVUs per FTE radiologist per year, with an average of 4,220 in academic groups and of 4,850 in nonacademic groups. (Approximately half their clinics had residents.) For 17 multispecialty clinics (largely the same institutions) in 1997, Conoley reported preliminary data showing an average of approximately 6,100 RVUs per FTE radiologist annually (Conoley P, personal communication, July 1999). These multispecialty clinics are a specialized environment with workloads that possibly are not representative of radiology groups in general; however, these data, like those from our reference survey (Table 1), suggest that our findings may be biased downward. According to the Conoley report, in 1997 the annual workload in RVUs per FTE radiologist among the 17 clinics varied by a factor of approximately two to one from the highest workload to the lowest; this was an indication of large variability, which is similar to what we found.

Cortegiano and James (7) reported an average of approximately 4,750 RVUs per FTE radiologist annually in a sample of 25 academic departments in 1996–1997. The calculations we performed with their data showed an average of 8,040 procedures per FTE radiologist and of 0.58 RVUs per procedure. For 1997–1998, Cortegiano (8) reported an average of approximately 5,200 RVUs per FTE radiologist annually in a sample of 32 (presumably largely overlapping) academic departments, which was an increase of approximately 9% in annual number of RVUs per FTE radiologist from that in the previous year. Our calculations with these data showed an average of approximately 8,300 procedures per FTE radiologist and of 0.60 RVUs per procedure. The two statistics for RVUs per procedure accord well with those from our survey (Table 1). The data on procedures and RVUs per FTE radiologist from these surveys are elevated relative to ours because they exclude from FTEs the time that department chairs spend in nonclinical work and the time that faculty members spend on externally funded research (but not other research); we did not make these exclusions. However, these adjustments affected the number of FTE radiologists on average by less than 10% (Cortegiano M, personal communication, July 1999), which is far less than the differ-

### TABLE 4

Results from Regression Analyses: Independent Variables

<table>
<thead>
<tr>
<th>Independent Variable</th>
<th>Coefficient</th>
<th>Standard Error</th>
<th>P</th>
<th>Coefficient</th>
<th>Standard Error</th>
<th>P</th>
<th>Coefficient</th>
<th>Standard Error</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>10,403</td>
<td>2,154</td>
<td>.001</td>
<td>2,268</td>
<td>1,226</td>
<td>.07</td>
<td>9.128</td>
<td>2,013</td>
<td>.001</td>
</tr>
<tr>
<td>Academic status†</td>
<td>−2,238</td>
<td>1,123</td>
<td>.05</td>
<td>−1,816</td>
<td>587</td>
<td>.003</td>
<td>−0.477</td>
<td>0.158</td>
<td>.003</td>
</tr>
<tr>
<td>Annual hours per FTE</td>
<td>0.425</td>
<td>1.119</td>
<td>.70</td>
<td>0.271</td>
<td>0.56</td>
<td>.63</td>
<td>−0.032</td>
<td>0.268</td>
<td>.91</td>
</tr>
<tr>
<td>Group size (FTE)</td>
<td>−59.7</td>
<td>32.5</td>
<td>.07</td>
<td>−30.0</td>
<td>15.5</td>
<td>.06</td>
<td>0.957</td>
<td>0.075</td>
<td>.001</td>
</tr>
<tr>
<td>Proportion of procedures that are “high productivity”</td>
<td>−4,782</td>
<td>11,975</td>
<td>.69</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Proportion of RVUs from “high productivity” procedures</td>
<td>6,355</td>
<td>1,717</td>
<td>.004</td>
<td>0.346</td>
<td>0.130</td>
<td>.009</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* In the logged regression analysis, all independent variables are also the natural logarithms, with the exception of academic status.
† Indicator variable: 1 if group is academic, 0 if group is nonacademic.
ence between their average workloads per FTE radiologist and ours. The higher values they found also suggest that our findings may be biased downward. For the 25 groups in 1996–1997 reported on by Cortegiano and James (7), we calculated the nonweighted SD of workload per FTE radiologist to be approximately 975 RVUs. For the 32 academic departments in 1997–1998, Cortegiano (8) reported an SD of 1,027 RVUs.

These SDs indicate substantial variability, which we also found, but the CV of the data for both years was 0.20, which is well below our 0.39 finding for academic groups (Table 2). These findings indicate that the workload among the academic groups in the above two studies was a good deal more homogeneous than that among the academic groups in our survey. For both years, the authors reported that there was no significant relationship between workload per FTE radiologist and department size, which is what we found.

The statistics on both the multispecialty clinics and the academic departments described above indicate that the workload measured in annual number of RVUs per FTE increased over time. However, the academic data cover a time span of only 1 year, and to our knowledge, neither group of researchers has reported on an explicit statistical test to assess whether the apparent increase in workload is real. The data from earlier surveys and analyses of Medicare claims that we reported also suggest an increase (1).

From personal communications with the authors, we understand that the surveys from which their data were gathered had excellent response rates—on the order of 80%. However, these rates were achieved in part by surveying only those practices that were likely to respond, and such groups may not be representative of all multispecialty or academic practices.

Analyzing data from our reference survey (1) points to an average workload of approximately 5,800 RVUs per FTE radiologist annually for all groups of all types, including the very small groups that were largely absent from the present survey. This is a somewhat weak estimate, because it combines a good estimate of procedures per FTE radiologist for all groups with data on average RVUs per procedure from Medicare claim analyses, which may not be typical of procedures performed in the non-Medicare population. However, as already noted, the present survey results regarding RVUs per procedure agree well with the findings regarding RVUs per procedure from the reference survey. The disparity is in the number of procedures, not in the number of RVUs per procedure, and it seems likely that the present survey results are biased downward.

Strengths and Limitations of the Study

Like all studies, the present study had strengths and limitations. The strengths included the relatively large number of groups in our study, the variety of types of groups in our survey, the acquisition and analysis of data on factors that were likely to affect workload, such as procedure mix and hours worked, the comparison of different measures of workload and their relative variability, and the multivariate analysis.

The principal weakness of our study was the poor survey response rate. Although our response rate was approximately twice that achieved by the Health Care Financing Administration, the federal government agency that oversees Medicare, in a survey that sought to obtain from physician practices financial information similar to that to which we sought, our response rate was low enough to make response bias a major concern. The paucity of responses from groups that contained two to four radiologists demonstrated that there was response bias related to group size. The data in Table 1, as well as the comparison of our findings with those from the other sources described earlier, suggest that there was also response bias in terms of workload. We have suggested a correction, but our estimate of the appropriate size of the correction must be recognized as imprecise. In future surveys, the response rate might be improved by seeking only workload data and omitting the detailed financial questions.

Another limitation of our study was that the data were group-wide aggregate values rather than information about individual radiologists. For some questions, especially those related to the workload of particular radiology specialty fields, individual-level data are needed. We have begun a research project to obtain and analyze such data.

In the survey, we did not ask about part-year sabbaticals or leaves, and this may have made the data on FTE radiologists slightly inaccurate, particularly for some academic groups. Our assumption of a 5-day work week may have introduced inaccuracies into the computation of the number of weeks radiologists worked annually, because we subtracted exactly 1/5 of a week for each day reported for vacation and other types of leaves. However, in the ACR’s 1995 Survey of Radiologists, a majority of full-time posttraining radiologists reported that they worked a 5-day week, and a majority of the remaining radiologists reported a 4.5- or 5.5-day week. Thus, the 5-day assumption should have introduced very small errors—on the order of 2% or less—into the total number of weeks worked annually. (Note that radiologists in a group may work 5 days weekly, even if the group provides 6- or 7-day services.)

Interpretation of Findings

All three measures of workload that we studied—annual number of procedures per FTE radiologist, annual number of physician work RVUs per FTE radiologist, and physician work RVUs per clinical work hour—had high variability. The variability remained high even when multiple factors that are likely to affect workload were considered simultaneously by using multivariate analysis. Predictions of the workload of individual groups that take into account all of these factors—that is, those made on the basis of the regression equations—still typically err by one-third relative to the actual values. In short, we observed substantial variability and had relatively little understanding of its source.

Hence, the average or median values should not be taken as norms. By doing so, one would be ignoring the variability or assuming that one understood it and that there was no good reason for it.

One possibly important source of variability that we did not study is that related to whether radiologists have to travel among facilities during a workday and whether the radiology department at a facility is entirely contiguous or split among floors or buildings. Another possibly important source of variability that we did not study is the extent to which groups, for reasons related to patient care quality or marketing, keep a radiologist on-site at a work location for more hours than are required solely to complete the work there (Conoley P, personal communication, July 1999).

Contrary to our expectations, workload measured in annual number of RVUs per FTE radiologist was as variable as workload measured in annual number of procedures per FTE radiologist (Table 2). In addition, the unexplained variability of workload measured in annual number of RVUs per FTE radiologist was not substantially lower than that of workload measured in annual number of procedures per FTE radiologist.
measured in annual number of procedures per FTE radiologist (see CV values in Table 3). Therefore, although RVUs are a conceptually superior metric, it is not clear whether in practice they offer an advantage to radiologists who want to know what workload is typical of that in groups like their own.

Although procedure mix is an important explainatory variable for annual number of RVUs but not for annual number of procedures, it is not clear whether this difference is an advantage or disadvantage of measuring workload in RVUs. It might be regarded as an advantage because procedure mix seems to be a priori an important source of variation, or it might be regarded as a disadvantage because additional information—namely that on procedure mix—is required to address the question “What is the average workload of a group like ours?”.

This is the appropriate question to ask if a group is seeking to compare its workload with typical workloads, and our research offers the following formula to answer it: Annual number of RVUs per FTE radiologist = 2,735 + 5,638 (proportion of RVUs from high productivity procedures) − 2,266 (if academic). This formula is the regression equation obtained when the nonsignificant variables (group size and hours worked) are omitted from the regression equation in the middle third of Tables 3 and 4. Omitting the nonsignificant variables is often considered to be appropriate.

Groups that use this formula need to recognize that the root mean squared error is approximately 4,650 and the CV is approximately one-third. Assuming a typical distribution, this implies that there is approximately a one-in-three chance that a group’s actual number of RVUs per FTE radiologist will diverge from the formula’s result by at least 1,650 and a similar one-in-three chance that the formula’s result will be at least 1.33 times the group’s actual workload or less than 0.67 times its actual workload. Again, the point is that variability is large, even after taking into account the factors that we found to affect workload, and the average values should not be taken as norms.

The results of multivariate analyses confirmed our expectation that workload in RVUs would be strongly affected by a group’s procedure mix and academic status. The relevant multivariate analysis results also showed that, as anticipated, the dominant factor in a group’s annual total number of RVUs was its size. Although expected, these findings should not be regarded as trivial, because the situation was very different with respect to the number of hours worked annually.

Our consistent finding of a lack of association between number of hours worked annually and annual workload, although totally unexpected, appears to be valid. Careful checking and rechecking revealed no data entry errors that could have accounted for this finding; and we cannot think of possible problems in the reporting of data on hours that could plausibly explain it. However, unless others confirm this finding, it should probably be regarded with some skepticism, because it remains difficult to understand how groups that work considerably different hours do not have at least vaguely similar differences in their annual workloads.

The recurrent, not quite statistically significant association that we found between larger group size and smaller workload per radiologist was probably a familiar statistical artifact technically known as “errors in variables” rather than a real effect. If such is the case, then it arose because the conceptually correct measure of group size was the average number of FTE radiologists during the 12-month period of the workload data, whereas in practice, group size is typically reported as of the survey date and therefore often differs somewhat from the conceptually correct measure.

In conclusion, taking into account probable response bias, our best estimate is that the average annual workload in 1995–1996 was approximately 6,000 physician work RVUs per FTE diagnostic radiologist in nonacademic groups and approximately 4,000 physician work RVUs per FTE diagnostic radiologist in academic groups; however, there is large, unexplained variance in every category of groups, so the average values should not be taken as norms. Although different procedures vary greatly in the amount of work each requires, group workload measured in annual number of RVUs per FTE radiologist unexpectedly proved to be as variable as workload measured in annual number of procedures per FTE radiologist. The results of multivariate analysis showed that procedure mix and academic status affect annual group workload measured in RVUs, as expected. However, for reasons that we do not understand, the annual number of clinical hours worked does not appear to affect workload. The factors we studied explain only about one-fourth of the variance in annual number of RVUs per FTE among groups. That three-fourths of the variance remains unexplained again points to the conclusion that average values should not be taken as norms.

Acknowledgments: We thank Michael Mabry, MA, for assisting in the development of the survey and providing expertise about CPT and RVUs; David Juba, MA, for extensive work on the survey; Patrick Conoley, MD, for providing data from multispecialty clinics and thoughtful observations; and Philip Creviston, PhD, and Jerry Bushée, PhD, for helpful advice and comments.

References
9. Health Care Financing Administration. Medicare program: revisions to payment policies and adjustments to the relative value units under the physician fee schedule, other Part B payment policies, and establishment of the clinical psychologist fee schedule for calendar year 1998. 211 Federal Register 59048–59260 (1997).