Gender and Utilization of Ancillary Services

Ashish K. Jha, MD, Gilad J. Kuperman, MD, PhD, Eve Rittenberg, MA, David W. Bates, MD, MSc

OBJECTIVE: To determine whether gender is associated with the use of ancillary services in hospitalized patients.

DESIGN: A retrospective study of laboratory and radiology tests ordered for medical and surgical inpatients over 16-month and 20-month periods, respectively. Obstetric patients were excluded.

MEASUREMENTS AND MAIN RESULTS: Number of clinical laboratory and radiology tests per admission, their associated charges, and total charges per admission were measured. In crude analyses, women had 16.5% fewer clinical laboratory tests ($p < .0001$) with 18.8% lower associated charges ($p < .0001$) and 24.4% fewer radiology tests ($p < .0001$) with 15.6% lower associated charges ($p < .0001$) than men. Total charges for the admission were lower for women in both the clinical laboratory study period (\$16,178 vs \$18,912, $p < .0001$) and the radiology study period (\$14,621 vs \$18,182, $p < .0001$). When adjusted for age, race, insurance status, service, diagnosis-related-group weight, and length of stay, these differences were smaller but persisted: women had 3.7% fewer laboratory tests performed ($p < .001$) with 4.8% lower associated charges ($p < .001$). In similarly adjusted analyses for radiology studies, women received 10.4% fewer radiology examinations ($p < .001$), with 4.1% lower associated charges ($p < .01$). There were no significant differences in the adjusted total charges in the laboratory group (\$17,450 vs \$17,655, $p = .20$) and only a marginally significant difference in the radiology group (\$16,278 vs \$16,498, $p = .05$). When we compared ancillary utilization within the five largest diagnosis-related groups, these differences persisted.

CONCLUSIONS: Men receive more ancillary services than women, even after adjusting for potential confounders.

KEY WORDS: ancillary services; gender; utilization.

J GEN INTERN MED 1998;13:476-481.

Men and women often receive different medical care. However, these differences are not always based on clinical characteristics.^{1,2} Studies comparing care for women and men have yielded varying results depending in part on the care setting and variables studied.³⁻¹²

Among outpatients, women generally have had more visits to physicians and receive more care per visit, including more laboratory tests.³ Hibbard et al. found that the largest differences between women and men in utilization rates occurred among patients who had symptoms associated with the mildest morbidities. They concluded that gender differences in utilization were due to differences in ways that physicians perceive and evaluate symptoms.² Cleary et al. suggested that women's higher outpatient utilization might be largely due to reported health differences.⁴

Among inpatients, many studies on gender and resource utilization have examined specific disease entities such as coronary disease and the use of interventional

procedures. Tobin et al. found that men were 10 times more likely to be referred for cardiac catheterization after an abnormal radionuclide test, and that women were more likely to have true cardiac symptoms dismissed as having a noncardiac etiology.⁶ Ayanian and Epstein studied coronary angiography and revascularization in coronary heart disease and found 15% to 45% higher rates for men.⁷ The SAVE trial found that although women reported worse anginal symptoms, they received fewer catheterizations and coronary artery bypass grafts.⁸

Differences in care have been demonstrated in non-cardiac areas as well. For instance, among those who fit the criteria for chronic dialysis, men are more likely than women to receive dialysis,⁹ while women on chronic dialysis are approximately 30% less likely to receive a cadaveric renal transplant than their male counterparts.¹⁰ Gender also has been shown to be an independent predictor for non-gender-specific preventive care, such as screening for lung cancer; after controlling for known risk factors, one study found that men had increased rates of screening sputum samples.¹¹

Some research has failed to demonstrate any effect of gender on resource use. For instance, McGann et al. studied men and women admitted for suspected transient ischemic attacks and found no difference between men and women in the use of computed tomography of the head, carotid and cardiac Doppler ultrasonography, or carotid angiography.¹³ This finding may reflect the relative standardization of care and a correspondingly lower level of discretion available to physicians in this clinical context.

Subtle differences in the ways physicians treat men and women may be most apparent in discretionary areas of care. Ancillary services represent such an area, given that studies have suggested that up to 50% of ancillary services performed on hospitalized patients may be unnecessary.¹⁴⁻¹⁷ In one study addressing whether women are treated differently from men, Bernard analyzed resource utilization of patients hospitalized on the medicine service and found that women had longer lengths of stay yet used fewer ancillary "relative value units."¹⁸

Because ancillary testing of inpatients is an area in which discretion plays a large role, and because earlier studies suggest differences in utilization by gender of several important inpatient services, we undertook a study

---

Received from the Division of General Medicine and Primary Care, Department of Medicine (AKJ, ER, DWB), and Center for Applied Medical Information Systems Research (GJK), Brigham and Women's Hospital, Harvard Medical School, Boston, Mass.

Address correspondence and reprint requests to Dr. Bates: Division of General Medicine and Primary Care, Brigham and Women's Hospital, 75 Francis St., Boston, MA 02115.

with the following goals: First, to compare crude differences between men and women in clinical laboratory and radiology tests performed for inpatients in the hospital, and then to determine the effect of gender on utilization after adjusting for a number of potential confounders.

METHODS

Patients

In the clinical laboratory group, we included all adults admitted to the medical and surgical services of Brigham and Women's Hospital, a 726-bed tertiary care hospital, over a 16-month period (February 1993 through May 1994). In the radiology group, we included all adults admitted to these services during a 20-month period (April 1993 through November 1994). Different periods were used because it took longer to map the charge data for all involved radiologic procedures to the involved orders. Obstetric patients were excluded for both the clinical laboratory and the radiology groups. Data on clinical laboratory and radiologic services used were obtained from the clinical database, while demographics and resource utilization information were obtained from the hospital's fiscal database. We excluded 612 (2.2%) of the clinical laboratory patient admissions and 444 (1.1%) of the radiology patient admissions because of incomplete data.

Information collected included age, gender, race, insurance status, clinical service at admission, discharge diagnosis-related group (DRG) and its associated weight, length of stay, and total charges. Insurance status was categorized as uninsured or Medicaid versus other (including, for example, Medicare, fee-for-service, and HMO). At the time of the study, this was the primary distinction driving resource utilization because there was relatively little capitation in our market. Weights are assigned by the Health Care Financing Administration to each DRG and are used to determine reimbursement for a specific DRG, so that they measure the expected intensity of resource use and, indirectly, severity of illness by diagnosis.⁵

Outcome Variables

Primary outcomes were laboratory tests, their associated charges per hospital admission, and number of radiology tests with their associated charges per admission. A test was defined as an ancillary service ordered by a physician that had an individual, separate charge code associated with it. We also looked at the total charges per admission to determine if any differences in the ancillary services were part of a larger trend that existed in other services used by hospitalized patients.

Analysis

The unit of analysis was the admission. Results of number of tests and associated charges are reported as

means with standard deviations as well as medians with 25% and 75% quartiles. Initial univariate comparisons were done using both the two-sample Student's *t* test and the Wilcoxon rank-sum statistic. Subsequently, we performed multivariate linear regression analyses adjusting for covariates known or suspected to affect ancillary service utilization. In the multivariate model, we adjusted for age, race, service at time of admission, primary insurance class, DRG weight, and length of stay. Because length of stay and DRG weight could be collinear, we looked for evidence of this using a variance inflation factor test, and found no evidence of significant collinearity for either the laboratory or radiology group.

Also, because the prevalence and quality of many clinical problems differ in men and women, case mix is a potential confounder. To address this issue, we examined the five most common DRGs in our hospital: major joint and limb procedure, lower extremity (DRG 209); percutaneous cardiovascular procedures (DRG 112); chest pain (DRG 143); major chest procedure (DRG 75); and heart failure and shock (DRG 127). The DRGs were evaluated individually and pooled using the same outcome variables mentioned above, adjusting for age, race, insurance category, and in the pooled analysis, adjusted for DRG. Although we would not expect statistically significant differences for individual DRGs because of small sample sizes, we felt it was important to evaluate the consistency of the effect within common DRGs because any differences present with groups with similar diagnoses would strengthen the overall inference that gender differences were not simply due to case mix. All analyses were done using SAS (Release 6.03, SAS Institute, Cary, NC, 1988).

RESULTS

Of the 28,055 admissions to the hospital in the clinical laboratory group, 14,101 (50.3%) were for women (Table 1). More men than women were white (76.9% vs 70.4%, $p < .0001$), but men and women had similar ages and insurance status. Compared with women, men had higher median DRG weights (1.19 vs 1.03, $p < .0001$) and longer lengths of stay (median [25%-75% quartiles]: 5 days [2-8 days] vs 4 days [2-8 days], $p < .0001$, Wilcoxon). The percentage of episodes that were readmissions in the clinical laboratory group was 22.1% for women and 23.9% for men; and in the radiology group, these figures were 22.8% for women and 24.9% for men.

The radiology group included 39,489 admissions, among which 21,107 (53.5%) were for women. Again, more men were white (76.3% vs 70.8%, $p < .0001$), but men and women had comparable age and insurance status. Men had higher median DRG weights (1.13 vs 1.00, $p < .0001$) and longer lengths of stay (median [25%-75% quartiles]: 4 days [2-7 days] for women, 4 days [2-8 days] for men, $p < .0001$, Wilcoxon).

During the study, women received 16.5% fewer clinical laboratory tests than men (48.0 ± 93.7 vs $57.5 \pm$

Table 1. Characteristics of Study Population

Characteristic	Laboratory Group		Radiology Group	
	Men (n = 13,954)	Women (n = 14,101)	Men (n = 18,382)	Women (n = 21,107)
Mean age $\pm$ SD, yr	54.7 $\pm$ 17.3	54.9 $\pm$ 18.4	54.0 $\pm$ 17.6	54.0 $\pm$ 18.2
Service, n (%)				
Medical	7,483 (53.6)	7,358 (52.2)	8,596 (46.8)	10,828 (51.3)
Surgical	6,471 (46.4)	6,743 (47.8)	9,786 (53.2)	10,279 (48.7)
DRG weight* (range)	1.19 (0.79–2.35)	1.03 (0.75–1.99)	1.13 (0.78–2.19)	1.00 (0.74–1.73)
White, n (%)	10,737 (76.9)	9,930 (70.4)	14,023 (76.3)	14,938 (70.8)
Medicare, HMO, or private insurance, n (%)	12,268 (87.9)	12,271 (87.0)	16,064 (87.4)	18,396 (87.2)
Length of stay, d (range)	5 (2–8)	4 (2–8)	4 (2–8)	4 (2–7)

*DRG indicates diagnosis-related group. Results are medians; ranges (in parentheses) are 25% and 75% quartiles.

113.7, $p < .0001$; Table 2). There was an 18.8% difference in associated laboratory charges (\$1,447 $\pm$ \$3,219 for women vs \$1,783 $\pm$ \$3,896 for men, $p < .0001$). The mean total charges per admission were \$16,178 $\pm$ \$26,718 for women and \$18,912 $\pm$ \$35,087 for men (-14.5% , $p < .0001$).

In the radiology study, women received 24.4% fewer tests (1.89 $\pm$ 3.53 vs 2.50 $\pm$ 4.23, $p < .0001$). Correspondingly, the associated charges for the radiology tests were 15.6% lower for women (\$286 $\pm$ \$540 vs \$339 $\pm$ \$603, $p < .0001$). In this cohort, the average total charges per admission were 19.6% smaller for women (\$14,621 $\pm$ \$24,726 vs \$18,182 $\pm$ \$33,499, $p < .0001$).

After adjusting for potential confounders including age, race, insurance status, DRG weight, and length of stay, we found that although the difference was smaller, women still received fewer laboratory tests per admission: -3.7% (51.8 $\pm$ 0.4 for women vs 53.8 $\pm$ 0.4 for men, $p < .001$, model multiple correlation coefficient [R^2] = .76; Table 3). In the adjusted associated laboratory charges, there was a comparable decrease in the size of the difference to -4.8% (\$1,576 $\pm$ \$17 for women vs \$1,656 $\pm$ \$17 for men, $p < .001$, model R^2 = .69). After adjusting for potential covariates, total charges per admission were slightly lower for women than for men, but the difference was not significant: -1.2% (\$17,450 $\pm$ \$113 for women vs \$17,655 $\pm$ \$113 for men, $p = .20$, model R^2 = .82). If

length of stay is excluded from these models, the strength of the association between gender and the outcomes decreases for outcomes (tests per admission, charges for tests, and total charges) and is no longer significant, suggesting that length of stay is an important confounder. We performed a test for collinearity between length of stay and DRG weight, and found no significant collinearity.

In adjusted analyses, using the number of radiology tests as the outcome, we also found a smaller difference than in crude analyses with women receiving 2.07 $\pm$ 0.02 radiology tests versus 2.31 $\pm$ 0.02 for men (-10.4% , $p < .0001$, model R^2 = .48). The adjusted radiology charges were \$305 $\pm$ \$3 for women and \$318 $\pm$ \$3 for men (-4.1% , $p < .01$, model R^2 = .38). In the adjusted analyses using total charges as the outcome for this period, women had modestly lower charges (\$16,218 $\pm$ \$96 for women vs \$16,498 $\pm$ \$102 for men, -1.7% , $p = .05$, R^2 = .78). For this group, we repeated the analyses without length of stay, and found that the magnitude of the association between gender and the outcomes again decreased, although the association between tests per admission remained significant ($p < .0001$).

When we studied the patients with the five most frequent DRGs, there was variation by DRG, but the utilization, adjusted for race, age, and insurance status, was the same or lower for women in both the clinical laboratory group (Table 4) and the radiology groups (Table 5) for all

Table 2. Tests and Charges per Admission*

Outcome	Laboratory Group			Radiology Group		
	Men (n = 13,954)	Women (n = 14,101)	Difference, [†] %	Men (n = 18,382)	Women (n = 21,107)	Difference, [†] %
Tests performed, n	57.5 $\pm$ 113.7 28 (14–60)	48.0 $\pm$ 93.7 24 (12–47)	-16.5	2.50 $\pm$ 4.24 1 (1–3)	1.89 $\pm$ 3.53 1 (0–2)	-24.4
Charges for tests performed, \$	1,783 $\pm$ 3,896 758 (348–1,780)	1,447 $\pm$ 3,219 636 (279–1,316)	-18.8	339 $\pm$ 603 100 (0–463)	286 $\pm$ 540 76 (0–365)	-15.6
Total charges, \$	18,912 $\pm$ 35,087 10,109 (5,174–21,066)	16,178 $\pm$ 26,718 8,848 (4,971–17,596)	-14.5	18,182 $\pm$ 33,499 9,454 (4,858–20,280)	14,621 $\pm$ 24,726 8,078 (4,695–15,661)	-19.6

*Results are reported as means $\pm$ SD and also as medians with 25th and 75th percentile in parenthesis.

[†]p Value $< .0001$ for all, using Student's t test, Wilcoxon.

Table 3. Multivariate Analyses: Tests per Admission and Charges per Admission*

Outcome	Laboratory Group				Radiology Group			
	Men, Mean $\pm$ SD (n = 13,954)	Women, Mean $\pm$ SD (n = 14,101)	Difference, %	R ²	Men, Mean $\pm$ SD (n = 18,328)	Women, Mean $\pm$ SD (n = 21,107)	Difference, %	R ²
Tests performed per admission, n	53.8 $\pm$ 0.4	51.8 $\pm$ 0.4	-3.7*	.76	2.31 $\pm$ 0.02	2.07 $\pm$ 0.02	-10.4*	.48
Charges for tests performed per admission, \$	1,656 $\pm$ 17	1,576 $\pm$ 17	-4.8*	.69	318 $\pm$ 3	305 $\pm$ 3	-4.1*	.38
Total charges per admission†, \$	17,655 $\pm$ 113	17,450 $\pm$ 113	-1.2‡	.82	16,498 $\pm$ 102	16,218 $\pm$ 96	-1.7§	.78

* $p < .0001$.

†Adjusted by age, diagnosis-related-group (DRG) weight, race, service, and insurance status.

‡ $p = .20$.§ $p = .05$.

DRGs but one. In pooled analyses, controlling for DRG, race, age, and insurance status, we found that women received 10% fewer laboratory tests (39.8 ± 1.3 vs 44.2 ± 1.3 , $p = .02$) with \$132 fewer charges (\$1,172 vs \$1,304, $p = .03$). In the similarly pooled and adjusted radiology patients, women received 2.51 ± 0.06 tests, while men received 2.82 ± 0.06 tests (-11%, $p = .0003$). The associated charges for these radiology tests were \$235 for women and \$263 for men (-11%, $p = .008$).

DISCUSSION

In the inpatient setting, we found that women received fewer clinical laboratory and radiology tests than

men did. On initial, unadjusted analysis, women received 9.5 fewer laboratory tests (charges, \$336) and 0.61 fewer radiology tests (charges, \$53) than did men. After adjusting for various demographic and clinical factors that may affect the rate of ancillary services use, the gender differences became considerably smaller but persisted; women received 2.0 fewer laboratory tests (charges, \$80) and 0.31 radiographs (charges, \$28). When we analyzed the pooled data from the five most common DRGs, controlling for DRG and other potential confounders, we found significant, large differences between men and women for both radiology and clinical laboratory tests. Finally, although we found trends toward smaller total charges for women, these differences were only marginally or not statistically significant.

Table 4. Mean Number of Laboratory Tests and Their Associated Charges in the Five Most Common Diagnosis-Related Groups

Diagnosis-Related Group	Men (Mean $\pm$ SD)	Women (Mean $\pm$ SD)	Difference, %	p Value
Major joint & limb procedure, lower extremity (DRG 209) (n = 1,070)				
Tests, n	33.3 $\pm$ 1.7	33.0 $\pm$ 1.3	-0.9	.90
Charges, \$	942 $\pm$ 59	921 $\pm$ 45	-2.2	.77
Percutaneous cardiovascular procedures (DRG 112) (n = 1,043)				
Tests, n	41.4 $\pm$ 1.7	38.0 $\pm$ 2.4	-8.2	.24
Charges, \$	1,113 $\pm$ 55	1,023 $\pm$ 76	-8.1	.34
Chest pain (DRG 143) (n = 931)				
Tests, n	20.4 $\pm$ 0.7	19.3 $\pm$ 0.7	-5.4	.25
Charges, \$	669 $\pm$ 21	624 $\pm$ 20	-6.7	.13
Major chest procedure (DRG 75) (n = 776)				
Tests, n	88.5 $\pm$ 5.7	68.0 $\pm$ 6.3	-23.2	.02
Charges, \$	2,804 $\pm$ 200	2,190 $\pm$ 224	-21.9	.05
Heart failure and shock (DRG 127) (n = 641)				
Tests, n	46.4 $\pm$ 2.5	45.9 $\pm$ 2.5	-1.1	.90
Charges, \$	1,323 $\pm$ 76	1,311 $\pm$ 74	-0.9	.91
Total five* (n = 4,461)				
Tests, n	44.2 $\pm$ 1.3	39.8 $\pm$ 1.3	-10.0	.02
Charges, \$	1,304 $\pm$ 43	1,172 $\pm$ 44	-10.1	.03

*Adjusted for age, race, insurance category, and DRG.

Table 5. Mean Number of Radiology Tests and Their Associated Charges in the Five Most Common Diagnosis-Related Groups

Diagnosis-Related Group	Men (Mean $\pm$ SD)	Women (Mean $\pm$ SD)	Difference, %	p Value
Major joint & limb procedure, lower extremity (DRG 209) (n = 1,070)				
Tests, n	1.42 $\pm$ 0.07	1.26 $\pm$ 0.05	-11.3	.07
Charges, \$	159 $\pm$ 9	133 $\pm$ 7	-16.4	.02
Percutaneous cardiovascular procedures (DRG 112) (n = 1,043)				
Tests, n	2.07 $\pm$ 0.06	2.05 $\pm$ 0.08	-1.0	.92
Charges, \$	97 $\pm$ 10	129 $\pm$ 13	24.8	.05
Chest pain (DRG 143) (n = 931)				
Tests, n	1.03 $\pm$ 0.02	1.00 $\pm$ 0.02	-2.9	.58
Charges, \$	126 $\pm$ 7	124 $\pm$ 7	-1.6	.83
Major chest procedure (DRG 75) (n = 776)				
Tests, n	7.98 $\pm$ 0.3	6.63 $\pm$ 0.3	-16.9	.002
Charges, \$	821 $\pm$ 34	667 $\pm$ 36	-18.8	.002
Heart failure and shock (DRG 127) (n = 641)				
Tests, n	1.87 $\pm$ 0.09	1.80 $\pm$ 0.09	-3.7	.62
Charges, \$	235 $\pm$ 16	228 $\pm$ 16	-3.0	.78
Total five* (n = 4,461)				
Tests, n	2.82 $\pm$ 0.06	2.51 $\pm$ 0.06	-11.0	.0003
Charges, \$	263 $\pm$ 7	235 $\pm$ 7	-10.6	.008

*Adjusted for age, race, insurance category, and DRG.

Our initial, crude analyses demonstrated large differences in ancillary utilization between men and women, but these differences shrank considerably when adjusted for covariates. This indicates that factors such as case mix are much more important as determinants of ancillary utilization than is gender. One of these determinants, the patient's length of stay, could have an associative as well as a causative relation to ancillary utilization, as performing more tests could be associated with longer length of stay owing to the sequelae of false-positive results. However, excluding length of stay from the models decreased the differences in utilization between men and women because men stayed longer than women.

We were concerned that the differences that persisted were due to differences in the disease processes that affect men and women, so we chose the five most common DRG categories between the two study groups, evaluated them individually, and also pooled them for further analysis. Only a few of the specific DRGs had statistically significant differences based on gender, probably because the sample size in each group lacked the power to detect small differences with consistency. However, when we pooled the data from the five DRGs, the differences in both the laboratory study and the radiology study were significant.

We found only modest differences in total charges for women. As ancillary services account for a small fraction of the total charges, we did not have the power to detect a modest difference in total charges if that difference was based on ancillary service use. The difference we found in total charges was approximately equal to the sum of dif-

ferences for radiology and clinical laboratory tests. The possibility of gender difference in ancillary test use but not in other areas of resource utilization is consistent with our hypothesis that different treatment based on the patient's gender might be more evident in areas of greater physician discretion.

The differences in clinical laboratory resource we found were clearly statistically significant, but small, particularly in the adjusted analyses. Thus, their clinical significance is debatable. However, the direction of the results (less resource use in women) is consistent with a large series of studies in other domains in which clinical significance is clear.^{7-10,19} The effect sizes we observed (3.7% for clinical laboratory tests and 10.4% for radiology tests) were smaller than those observed in some other domains such as cardiac procedures, in which rates up to 45% higher have been reported.⁸

There are several possible explanations for the gender differences we observed. There could be clinical differences between men and women that we were not able to account for in the adjusted analyses. In particular, men could be sicker than women. However, we did adjust for DRG weight and length of stay, both of which are measures of case mix. Another possibility is that men request more tests than women. However, this is unlikely to be the case among inpatients, whose wishes typically play a smaller role than those of patients in other settings.

Another issue that most likely plays a role is gender differences in test ordering on the part of physicians. Several studies have evaluated the effect of physician gender

on test utilization.^{1,2,12,20} Lurie et al. found that women were more likely to receive Pap smears and mammograms if they saw female physicians,²⁰ possibly representing better care. However, in another study, Rosen et al. found that patients seeing women physicians were more likely to receive radiographs than men,¹² and this effect was stronger for male patients than for female patients. Because we did not know the gender of the physicians ordering the tests in this study, we could not explore these issues. However, male and female patients will continue to see both male and female physicians, and in our view the goal should be to develop strategies to ensure that patients receive appropriate care independent of their gender or their physician's.

From our data, we cannot say which group received more appropriate care. However, both overtutilization and underutilization represent suboptimal care. Data about whether most variation is due to overutilization or underutilization are sparse, but in one study, Leape et al. found that inappropriate use accounted for only a small fraction of variance in the use of several interventional procedures.²¹

This study had several limitations. Our methods of case-mix adjustment, using DRG weights and length of stay, are relatively crude. However, the overall model R^2 was high, suggesting that the combination of DRG weights and length of stay accounted for much of the variability in ancillary testing. Also, we could not examine whether the differences in ancillary services utilization had any effect on outcome; therefore, we could not address the issue of whether men or women received more appropriate care.

We found gender differences for inpatients in the use of ancillary services, with women receiving fewer tests. However, we also found large variation between different DRGs, and most of the gender differences are accounted for by known predictors of ancillary utilization such as length of stay and DRG weight. The differences based on gender are persistent but small and may not exist in other areas of resource use. These findings are consistent with previous reports and the hypothesis that in areas of high physician discretion, factors such as patient's gender may play a larger, possibly inappropriate role in resource allotment.

This study was supported by research grant R01-HS07107-01 from the Agency for Health Care Policy and Research. Dr. Jha was supported in part by the Carl W. Walter fund of Harvard Medical School.

The authors thank Anthony L. Komaroff, MD, for comments on a previous version of this manuscript.

REFERENCES

1. Lerner DJ, Kannel WB. Patterns of coronary heart disease morbidity and mortality in the sexes: a 26-year follow-up of the Framingham population. *Am Heart J*. 1986;111:383-90.
2. Hibbard JH, Pope CR. Another look at sex differences in the use of medical care: illness orientation and the types of morbidities for which services are used. *Women Health*. 1986;11:21-36.
3. Verbrugge LM, Steiner RP. Physician treatment of men and women patients: sex bias or appropriate care? *Med Care*. 1981;19:609-32.
4. Cleary PD, Mechanic D, Greenley JR. Sex differences in medical care utilization: an empirical investigation. *J Health Soc Behav*. 1982;23:106-19.
5. Lauer MS, Pashkow FJ, Snader CE, Harvey SA, Thomas JD, Marwick TH. Gender and referral for coronary angiography after treadmill testing. *Am J Cardiol*. 1996;78:278-83.
6. Tobin JN, Wassertheil-Smoller S, Wexler JP, et al. Sex bias in considering coronary bypass surgery. *Ann Intern Med*. 1987;107: 19-25.
7. Ayanian JZ, Epstein AM. Differences in the use of procedures between women and men hospitalized for coronary heart disease. *N Engl J Med*. 1991;325:221-5.
8. Steingart RM, Packer M, Hamm P, et al. Sex differences in the management of coronary artery disease. Survival and Ventricular Enlargement Investigators. *N Engl J Med*. 1991;325:226-30.
9. Kjellstrand CM, Logan GM. Racial, sexual and age inequalities in chronic dialysis. *Nephron*. 1987;45:257-63.
10. Held PJ, Pauly MV, Bovbjerg RR, Newmann J, Salvatierra O Jr. Access to kidney transplantation. Has the United States eliminated income and racial differences? *Arch Intern Med*. 1998;148: 2594-600.
11. Franks P, Clancy CM. Physician gender bias in clinical decision-making: screening for cancer in primary care. *Med Care*. 1993;31:213-8.
12. Rosen MP, Davis RB, Lesky LG. Utilization of outpatient diagnostic imaging. Does the physician's gender play a role? *J Gen Intern Med*. 1997;12:407-11.
13. McGann KP, Marion GS, Szweczyk MB, Davis SW. Absence of sex differences in the evaluation of patients hospitalized for transient ischemic attacks. *J Fam Pract*. 1994;39:134-139.
14. Eisenberg JM, Williams SV, Garner L, Viale R, Smits H. Computer-based audit to detect and correct overutilization of laboratory tests. *Med Care*. 1977;15:915-21.
15. Williams SV, Eisenberg JM. A controlled trial to decrease the unnecessary use of diagnostic tests. *J Gen Intern Med*. 1986;1:8-13.
16. Schroeder SA, Kenders K, Cooper JK, Piemme TE. Use of laboratory tests and pharmaceuticals. Variation among physicians and effect of cost audit on subsequent use. *JAMA*. 1973;225:969-73.
17. Dixon RH, Laszlo J. Utilization of clinical chemistry services by medical house staff. An analysis. *Arch Intern Med*. 1974;134: 1064-7.
18. Bernard AM, Hayward RA, Rosevear JS, McMahon LF. Gender and hospital resource use. Unexpected differences. *Eval Health Prof*. 1993;16:177-89.
19. Khan SS, Nessim S, Gray R, Czer LS, Chaux A, Matloff J. Increased mortality of women in coronary artery bypass surgery: evidence for referral bias. *Ann Intern Med*. 1990;112:561-7.
20. Lurie N, Slater J, McGovern P, Ekstrum J, Quam L, Margolis K. Preventive care for women. Does the sex of the physician matter? *N Engl J Med*. 1993;329:478-82.
21. Leape LL, Park RE, Solomon DH, Chassin MR, Koseoff J, Brook RH. Does inappropriate use explain small-area variations in the use of health care services? *JAMA*. 1990;263:669-72.