

Pay-for-Performance—Financial Health Disparities and the Impact on Healthcare Disparities

Rodney G. Hood, MD

Key words: health insurance ■ race/ethnicity

© 2007. From W. Montague Cobb/NMA Health Institute (advisor), Multicultural IPA (president) and Careview Medical Group (CEO), San Diego, CA. Send correspondence and reprint requests for *J Natl Med Assoc.* 2007;99:xxx-xxx to: Dr. Rodney G. Hood, CEO, Careview Medical Group, 292 Euclid Ave., Suite 210, San Diego, CA 92114; phone: (619) 262-7523; e-mail: rghood@cox.net

Payment reform has recently been introduced that will reimburse providers based upon performance goals that are expected to become a significant component of future provider reimbursement formulas. This is highlighted by the comment of Mark McClelland, previous Centers for Medicare and Medicaid Services (CMS) administrator, when he predicted,

In the next 5–10 years pay-for-performance-based compensation could account for 20% to 30% of what Medicare pays providers.

The profession of medicine is highly controlled and influenced by governmental and private health policy sectors that initiate policies that create a health system governed by the “medical Holy Trinity” consisting of medicine, policy and finance. Since health polices and financial formulas are the main tools used to regulate healthcare providers and consumers, we must be concerned that the drive to contain costs is not emphasized at the expense of quality improvement. If the new payment reform is to achieve ultimate quality improvement with better health outcomes such as longer life expectancy and improved quality of life then we must develop better *outcome* quality measures rather than the commonly used *process* measures. The policies that assume quality process measures will improve quality and better health outcomes have no strong evidence-based data to support this conclusion.¹

Significant payment reform has been introduced that will greatly impact the quality of medical care received by all Americans—but have any health policy studies involved a comprehensive analysis evaluating the impact of

current and future payment reforms on the populations with the highest disease burdens and greatest healthcare and health disparities? A search of the medical literature suggest the answer to this question is no. As the representative for the National Medical Association (NMA) in the American Medical Association (AMA) House of Delegates during the AMA forum on pay-for-performance (P4P), I testified on December 5, 2004 that

Pay-for-Performance programs that do not consider specific health disparities risk variables such as socioeconomic status, geographic location, race, ethnicity and level of disease burdens can create the real potential to economically penalize and cause unintended disincentives for individual physicians, medical groups and health institutions that have traditionally provided health services for these high-risk populations. These inequities will further worsen quality of care in high-risk populations and worsen health-care disparities.

If the goal of P4P is to apply the best evidence-based medicine to improve medical quality in a cost-efficient manner, then we must ask: Whose evidence? Based upon what assumptions? Improved quality for whom? And at what costs? Racial and ethnic health disparities are real, and these disparities vary depending upon the ethnic group identified. Multiple studies have shown that utilization of therapeutic and diagnostic procedures such as coronary artery bypass graft (CABG), angioplasty, endarterectomies, knee and hip replacements, organ transplants and physical therapy found blacks have the greatest disease burdens and widest health disparities but received these procedures less.^{2,3} However, when we look at organ or limb removal such as orchiectomy, limb amputation and hysterectomies, blacks are less likely to choose these options but received these procedures more frequently than whites.^{2,3} This poses that when the therapeutic modality suggests more is better, blacks receive less, but when the therapeutic modality suggests less is better, blacks receive more when compared to whites.

The Agency for Healthcare Research and Quality (AHRQ) Annual National Healthcare Disparities Report (2004) found that all major minority groups received lower quality of medical care and poorer access to care when compared to whites.⁴ Blacks, Latinos, American Indians and Asians were found to receive 66%, 50%, 33% and 10% less quality of care, respectively, compared to whites. These same ethnic groups were found to have 30–90% less access to care compared to whites, with Latinos having the highest difficulty with access. These findings were 60–80% worse if the patient was poor. Among Medicare beneficiaries enrolled in managed care plans, African Americans compared to whites received poorer quality of care utilizing the NCQA Health Plan Employer Data and Information Set (HEDIS) for thousands of patients enrolled in Medicare health plans looking at: 1) breast screening mammography, 2) use of beta-blockers post-heart attack, 3) eye examinations for diabetics, and 4) follow-up visits after psychiatric hospitalizations, as reported by Schneider in *JAMA* in 2002.⁵ The Institute of Medicine report “Unequal Treatment: Confronting Racial and Ethnic Disparities in Health Care” convincingly confirmed the reality and some of the causes for racial and ethnic health disparities.³

The California Integrated Health Association (IHA) P4P initiative is commonly sighted as an example of a successful program in initiating P4P.⁶ The IHA began in July 2000 with a high-level working group of California healthcare leaders from health plans, physicians and medical directors who met to discuss a new statewide initiative for P4P. In January 2002, six California health plans (Aetna, Blue Cross, Blue Shield, CIGNA, HealthNet and PacifiCare) launched this new initiative. A score card of common performance measures were agreed upon, with clinical measures weighted at 50%, patient satisfaction weighted at 40% and information technology (IT) at 10%. Updates of this initiative increased IT to 20% and changed the clinical measures to 40%. The clinical measures included breast and cervical cancer screening, childhood immunizations, chlamydia screening, diabetes A1C measure, cholesterol–LDL measure and use of asthma medication.

I will share insights and experiences with participation in the California IHA P4P initiative in San Diego County, CA and its effects on the insured minority populations enrolled in commercial products that do not include Medicaid. The Multicultural Primary Care Physician Medical Group, Independent Physicians Association (MCIPA) was a participating medical group in 2003–2004. MCIPA is a for-profit organization of physicians that was established in San Diego County and was managed by the University of California–San Diego (UCSD) Health Network beginning in 1994. Since 2003, MCIPA has been managed by SynerMed Inc., a medical management company located in Los Angeles. MCIPA generates \$6 million yearly from commercial, senior and Medicaid direct health plan con-

tracts, and the group is composed of approximately 50 primary care physicians (PCPs) and >50 specialty healthcare providers. In 2003, the MCIPA had approximately 12,000 enrollees (8,000 commercial) with providers and enrollees that are ethnically diverse. Enrollees are mostly Latino and African American but include Asian, African and other immigrants and those of European descent. MCIPA providers and enrollees are predominantly located in the central and south regions of San Diego County. The San Diego County population is approximately 3 million, with 8,700 physicians, making a physician:population ratio in San Diego County at 1:350. However, the physician ratio for MCIPA service areas in the central and south regions of San Diego County, where the majority of MCIPA physicians are located, is approximately 1:1,200. Therefore, MCIPA service areas have a physician shortage of three times fewer physicians than other parts of the county. According to the San Diego County Health Needs Assessment Report (2004), populations with the highest disease burdens and greatest obstacles to access healthcare are found in the central and south regions, with African Americans suffering the highest disease burdens and Latinos the worst access; and populations living in the central and south regions of San Diego County have the highest hospitalization and death rates from diabetes, asthma, cardiovascular heart disease and cancer.⁷

Within the IHA initiative, the MCIPA's overall quality clinical and patient satisfaction ratings were scored as fair. However, were these ratings based upon accurate evidence? The quality indicator criteria utilized in the IHA model were based on aggregate data from clinical process measures that favored the majority populations and, therefore, physicians were disadvantaged if they served populations that were disproportionately minority with high disease burdens. The clinical process measures utilized in this initiative have known baseline disparities that are generally lower when minority populations are compared to the majority population. The UCLA Center for Health Policy Research conducted surveys and reported that the self-reported mammography and Pap screening rates for blacks and Latinos were similar to those of whites.⁸ However, multiple studies point out that self-reported cancer screening rates were much lower and less than whites when documented by medical records.^{9–14}

In 2000–2001, the overall national biennial breast screening rates for women ≥ 40 years were 50.6% for non-Hispanic white women, 40.5% for black women, 34.7% for Asian-American women, 36.3% for Hispanic women and 12.5% for native-American women; thus, resulting in 20–75% lower rates for minorities compared to whites.¹² These baseline population specific quality measure disparities are prevalent throughout the United States and exist for most other quality measures used in P4P. In 2003, the MCIPA, serving a disproportionate high minority population, recorded clinical quality measure scores of 54% and 50% for breast and cervical can-

cer screening, respectively.¹⁵ The medical groups in San Diego County scored a range of 40–86% for cervical screening and 54–86% for breast screening. The MCIPA provider–patient satisfaction sample scores for doctors working as a team, having helpful office staff, and visits starting on time were 89%, 86% and 34%, respectively. The MCIPA performed fairly well with patient satisfaction scores except for the “visits start on time,” for which MCIPA providers scored one of the lowest in the county.¹⁵ Many of the MCIPA physicians practice in low-income and physician-shortage areas that produce high patient loads. Physicians practicing in areas where many physicians choose not to practice are being penalized for serving underserved populations.

The reasons for providers’ low-quality performance scores with disproportionate enrollment of high-risk minority populations include multiple inequities with the data used for comparison and interpretation. The inequities encountered include: 1) inadequate baseline reimbursement; 2) administrative costs; 3) racial quality indicator disparities; 4) incomplete encounter data collection; 5) unfair quality indicator comparisons; 6) tiered physician networks and physician economic profiling; 7) de facto racial, ethnic and socioeconomic status (SES) discrimination; 8) geographic physician shortages, and the ultimate inequity; and 9) worsening of health disparities.

Inequity #1—Inadequate Baseline Reimbursement

In a capitated model, physicians’ health services are reimbursed based upon average costs with the assumption that the enrolled population has a bell-shaped curve “risk” distribution with an equal balance of low- and high-risk populations (Figure 1). If the enrolled population has an adverse high-risk selection based upon risk variables such as race, ethnicity, geographic location or SES, the average service costs are expected to be higher. If the medical group’s adverse high-risk popula-

tion is reimbursed at the lower rates of the average-risk population, the group will receive less compensation for the high-risk population served (Figure 2). In general, the contracting physicians or medical groups are reimbursed at the average-risk costs minus HMO administrative withholds, then reimbursement is more or less depending upon the number of services contracted and the group’s negotiating strengths or weaknesses. Therefore, a medical group with a disproportionate high-risk population enrollment and a weak negotiation position due to small enrollment will likely receive a rate between the low-risk versus average-risk rates.

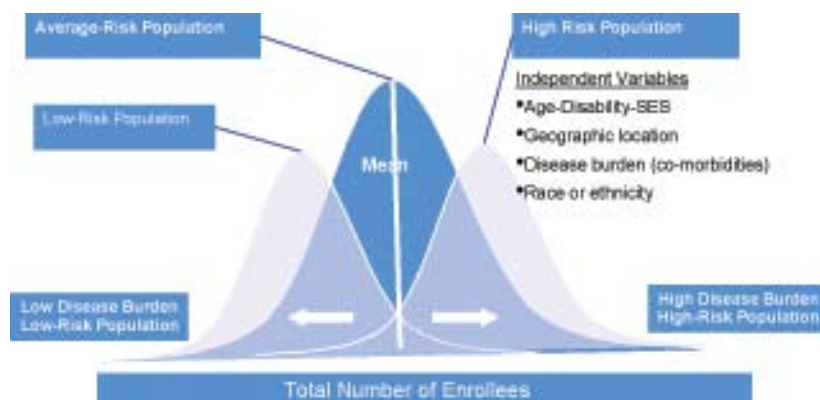
Inequity #2—Administrative Costs

The HMO withholds up to \$3–\$4 per member per month (pmpm) from participating physician groups to cover P4P incentive costs and is *not* extra money. The physician group’s P4P quality improvement program cost the medical group approximately \$1 pmpm to implement. A separate fee is charge to the medical group (\$2,000 for a small group) to cover costs of the patient survey. Therefore, the group incentive withholds, the group program costs, plus other fees further diminishes physicians’ reimbursements.

Inequity # 3—Racial Quality Indicator Disparities

Medical groups serving populations having the highest health disparities and the greatest disease burdens, such as blacks, Latinos and Asians, have lower average baseline quality indicator levels when compared to whites.⁹⁻¹⁴ When P4P quality indicator criteria are based upon low-risk groups, it creates disparities with quality measure goals that are disproportionately higher when compared to the high-risk groups. Therefore, groups serving populations of high disease burden (high risk) will receive little or no financial benefit from P4P incentive withholds and in fact can be penalized with even less reimbursement.

Figure 1. Supplemental material



Population and geographic variables, disease burdens and the assumed risk distribution in a capitated managed care model in California

Inequity #4—Incomplete Encounter Data Collection

The provider services encounter data are utilized to measure physician groups' levels of compliance with quality improvement goals. Clinical measures that are actually performed by the provider or performed outside the network but not recorded with encounter data collected by the health plan will not be captured. Physicians with less IT capacity tend to submit incomplete encounter data at higher rates. Therefore, incomplete collection of encounter data results in lower-quality indicator scores even when the clinical measure was accomplished.

Inequity #5—Unfair Quality Measure Comparisons

Each physician group's quality data are published as a quality report card that presumes to represent a quality-of-care comparison. Physicians serving disproportionate high-risk populations with baseline lower-quality indicators will be perceived as giving poor quality and therefore negatively affect enrollment. Incentives to access low-cost IT such as electronic medical records can enhance practice efficiencies and improve clinical quality reporting.

Inequity # 6—Tiered Physician Networks and Physician Economic Profiling

Tiered physician networks are composed of physicians or medical groups that are partitioned into different tiers based upon their cost efficiencies as deemed by the health plan. Utilizing economic profiling, selected physician groups that are deemed cost efficient are placed into a select network tier that offers patients lower copays and a more enriched benefit plan. Physicians or groups with high-risk populations based upon variables that include SES, geographic location, high disease burdens or comorbidities, and race/ethnicity are deemed less cost efficient and further penalized by lower-tiered plans that offer

higher copays, fewer benefits and encourage lower-risk patients not to enroll with traditional high-risk providers.

Inequity # 7—De Facto Racial, Ethnic and SES Discrimination

P4P as currently being implemented creates disincentives for physicians and medical groups to not enroll high-risk patients that are disproportionately ethnic minorities.¹⁶ This scenario encourages de facto racial, ethnic, social and economic discrimination. Thus, high-risk patients default to traditional healthcare providers, further worsening quality indicator data due to high disease burdens for high-risk populations.

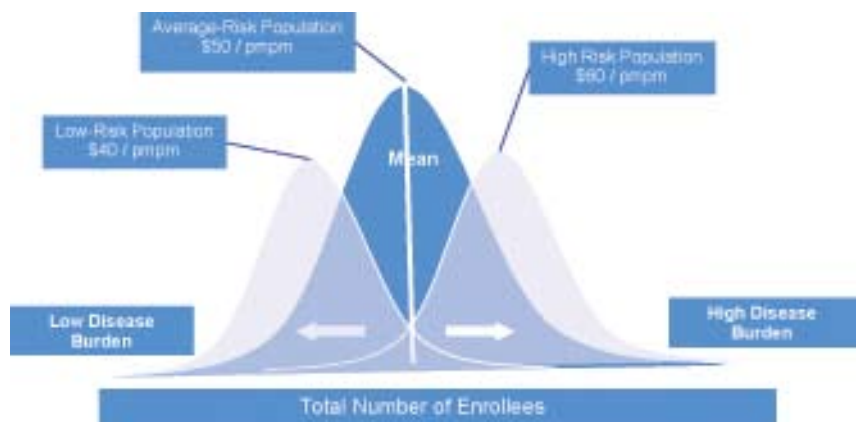
Inequity # 8—Geographic Physician Shortages

Many rural and urban communities have significant physician shortages, and it is not uncommon for these areas to correlate with populations that have the greatest medical needs. Physicians who work in underserved communities are many times forced to accommodate heavier patient volumes, making it more difficult to avoid longer waits during office visits. Rather than penalize physicians who serve underserved populations with limited access, appropriate reimbursement and other resource incentives should be offered to increase providers' desire to practice in these underserved physician shortage communities.

Inequity # 9—Worsening of Health Disparities

P4P programs that do not equitably compensate for high-risk populations and utilize inaccurate quality indicator comparisons will not enhance the elimination of health disparities but may actually worsen health disparities. In New York, the publishing of a report card for physicians performing CABG surgery resulted in worsening of racial and ethnic disparities compared to whites (Figure 3).¹⁵ The percent of surgical disparities for coronary patients before

Figure 2. Population and geographic disease burdens and risk distribution utilized in managed care reimbursement formulas with estimated professional capitated costs [(\$) pmpm]



the report card for blacks and Latinos were 2.7 and 0.7, respectively; after the report card was published, the disparities increased to 5.0 for blacks and 3.2 for Latinos (Figure 3).¹⁶ Providers' desire to improve their report card scores resulted in an increase in healthier patients undergoing CABG surgery and fewer sick patients.

New York and Pennsylvania CABG report cards caused "cherry-picking." The publication of the report cards led to higher cost for both healthier patients, who got more CABG surgeries, and sicker patients despite having stable-to-declining surgery rates. The report cards led to roughly unchanged outcomes for healthy and much worse health outcomes for sick patients.¹⁷

LESSONS AND RECOMMENDATIONS

- Healthcare disparities are quality issues that came about because of healthcare inequities.**
 Recommendation: Cautiously adopt the concept of P4P as a tool to correct health disparities as a quality issue by addressing the structural health inequities.
- Baseline reimbursements should reflect the population's risk levels.**
 Recommendation: We need core payment reform. Reimbursement must reflect the population's level of risk, utilizing variables of race/ethnicity, disease burdens, geographic location and SES, plus P4P incentive payments should be based upon percent improvement of the actual groups' baseline quality measures rather than set levels that are based upon lower-risk populations.
- P4P is a potential tool to monitor and improve health disparities.**
 Recommendation: P4P has the potential to worsen health disparities. We must not allow the health

system to ignore or minimize the inequities of performance criteria. All performance measures must address risk factors such as population disease burdens, access disparities, geographic and SES disparities as well as race and ethnicity as independent health risk variables.

- Physician groups associated with larger networks perform better probably because of access to better management tools.**
 Recommendation: Medical practice integration and embracing IT will be imperative for success. Independent physicians and small physician groups must find ways to integrate their practices with larger entities in order to take advantage of cost efficiencies and access to IT.
- Health policy advocates should prioritize to bring about programs and legislation at both the state and national levels that promote these changes.**
 Recommendations: 1) Obtain grants to establish P4P quality improvement programs with physicians and medical groups serving high-risk populations. 2) Advocate for state health agencies and CMS to mandate core payment reform. HMOs and government payors must establish P4P reimbursement formulas that will ensure appropriate reimbursement rates for high-risk populations.

DISCUSSION

Rosenthal reported in *JAMA* (2005) an early experience with pay-for-performance.¹⁸ The study compared the P4P IHA initiative in California and followed three measures—cervical cancer screen, mammography and hemoglobin A1C. For all three measures, physician groups with baseline performance at or above the per-

Figure 3. Quality report card and worsening health disparities



Werner, et al. New York CABG Report Card 1991, *Circulation*

formance threshold for receipt of a bonus improved the least but garnered the largest share of the bonus payments (\$3.4 million). The study concluded:

Paying clinicians to reach a common, fixed performance target may produce little gain in quality for the money spent and will largely reward those with higher performance at baseline.

This conclusion supports the California IHA experience of the MCIPA.

An example of a quality issue that has gone unnoticed by the quality performance wonks can be highlighted by the African-American Heart Failure Trial (AHeFT).¹⁹ Over a mean of 10 months, African Americans with congestive heart failure (CHF) who were given isosorbide dinitrate-hydralazine (ISDN-H) combination experienced a 43% drop in relative mortality risk, a 39% decrease in the relative risk of heart failure hospitalization, and improved quality of life. After a year of being approved by the FDA registry, data suggest that no more than 20% of the target population is taking ISDN-H combination or its separate generic components. A study in 2005 reported in *Circulation* showed that the use of the ISDN-H combination was cost effective mainly due to decrease hospitalizations.²⁰

The use of ISDN-H for CHF in African Americans appears to meet criteria for a population-specific P4P quality improvement measure. The P4P equation = evidence-based medicine + cost-efficiency + patient centered healthcare. With the use of ISDN-H in African Americans, we have met P4P criteria seen with the P4P equation = decreased mortality (A-HeFT) + decreased hospitalizations + improved patient quality of life. This evidence creates the opportunity to create a health disparity quality improvement measure that will monitor the use of ISDN-H combination in the affected population.

CONCLUSION

The inconvenient truth is we must first address and eliminate the structural inequities that are built into the reimbursement and quality improvement models before we can expect to see true “quality improvement” that will achieve the goal of eliminating health disparities. P4P has the potential to be one of many tools that can improve quality and diminish health disparities. We must be proactive rather than prospective by identifying and integrating population-specific quality improvement measures when implementing any new healthcare reform. We can only solve the problem of eliminating health disparities by first eliminating the inequities that perpetuate these disparities. How?, you may ask. The answers are simple, but the will to implement needed solutions is lacking. We must find the collective will to re-engineer the health system to administer high-quality healthcare to all of America’s diverse populations. We should utilize the best evidence for quality measures that incorporates population-specific data that more accurately

represents all of America’s diverse populations and mandate a redistribution of resources through core payment reform that will more fairly compensate providers serving populations with disproportionate heavy disease burdens. Our resolve to change a health system that rewards the least needy with the greatest resources will be our greatest challenge.

REFERENCES

1. Werner RM, Bradlow ET. Relationship between Medicare's hospital compare performance measures and mortality rates. *J Am Med Assoc.* 2006;296:2694-2702.
2. Mayberry RM, Mili F, Vaid IGM, et al. Racial and ethnic differences in access to medical care: a synthesis of the literature. Prepared by the Morehouse Medical Treatment and Effectiveness Center (MMEDTEC). Menlo Park, CA: Henry J Kaiser Family Foundation; 1999.
3. Smedly BD, Stith AY, Nelson AR, eds. *Unequal Treatment: Confronting Racial and Ethnic Disparities in Health Care.* Washington, DC: National Academies Press; 2003.
4. The 2004 Annual National Healthcare Disparities Report: Quality of Care and Access to Care Comparisons by Selected Racial Groups 2000-2001. Rockville, MD: Agency for Healthcare Research and Quality (AHRQ), U.S. Dept. Health and Human Services. AHRQ Publication No 05-0014; December 2004.
5. Schneider EC, Zaslavsky AM, Epstein AM. Racial disparities in the quality of care for enrollees in Medicare managed care. *J Am Med Assoc.* 2002;287:1288-1294.
6. California Integrated Healthcare Association (IHA), Oakland, CA. www.ih.org. Accessed 03/15/07.
7. The 2004 Charting the Course IV: a San Diego County Health Needs Assessment. Community Health Improvement Partners (CHIP). www.sdchip.org. Accessed 03/01/05.
8. Ponce NA, Babey SH, Etzioni DA, et al. Cancer Screening in California: Findings from the 2001 California Health Interview Survey. UCLA Center for Health Policy Research; December 2003.
9. Champion VL, Menon U, McQuillen DH, et al. Validity of self-reported mammography in low-income African American women. *Am J Prev Med.* February 1998.
10. Goel MS, Wee CC, McCarthy EP, et al. Racial and ethnic disparities in cancer screening: the importance of foreign birth as a barrier to care. *J Gen Intern Med.* 2003;18:1028-1035.
11. Haitt RA, Pasick RJ, Stewart S, et al. Community-based cancer screening for underserved women: design and baseline findings from the breast and cervical cancer intervention study. *Prev Med.* 2001;33:190-203.
12. Kagay CR, Quale C, Smith-Bindman R. Screening mammography in the American elderly. *Am J Prev Med.* August 2006.
13. Maxwell AE, Bastani R, Warda US. Breast cancer screening and related attitudes among Filipino-American women. *Cancer Epidemiol Biomarkers Prev.* 1997;6(9).
14. Paskett ED, Tatum CM, Mack DW, et al. Validation of self-reported breast and cervical screening test among low-income minority women. *Cancer Epidemiol Biomarkers Prev.* 1996;5(9).
15. The 2005 California HMO Report Card, Office of Patient Advocate, Managed Care Research and Reports. www.opa.ca.gov/reports. Accessed 03/15/07.
16. Warner RM, Arch DA, Polsky D. Racial Profiling: the unintended consequences of coronary artery bypass graft report cards. *Circulation.* 2005;111:1257-1263.
17. Dranove D, Kessler D, McClellan M, et al. Is more information better? The effects of “Report Cards” on health care providers. *J Political Economy.* 2003;111(3).
18. Rosenthal MB, Frank RG, Zhonghe Li MA, et al. Early experience with pay-for-performance. *JAMA.* 2005;294:1788-1793.
19. Taylor AL, Ziesche S, Yancy C, et al. Combination of isosorbide dinitrate and hydralazine in blacks with heart failure. *N Engl J Med.* 2004;351:2049-2057.
20. Angus DC, Linde-Zwirble WT, Tam SW, et al. Cost-effectiveness of fixed-dose combination of isosorbide dinitrate and hydralazine therapy for blacks with heart failure. *Circulation.* 2005;112:3745-3753. ■