

© Health Research and Educational Trust

DOI: 10.1111/j.1475-6773.2009.01022.x

RESEARCH ARTICLE

Physicians and Coronary Heart Disease

Do Physicians Attend to Base Rates? Prevalence Data and Statistical Discrimination in the Diagnosis of Coronary Heart Disease

Nancy N. Maserejian, Karen E. Lutfey, and John B. McKinlay

Objective. To examine whether physicians attend to gender prevalence data in diagnostic decision making for coronary heart disease (CHD) and to test the hypothesis that previously reported gender differences in CHD diagnostic certainty are due to discrimination arising from reliance on prevalence data (“statistical discrimination”).

Data Sources. A vignette-based experiment of 256 randomly sampled primary care physicians conducted from 2006 to 2007.

Study Design. Factorial experiment. Physicians observed patient presentations of cardinal CHD symptoms, standardized across design factors (gender, race, age, socioeconomic status).

Data Collection. Structured interview.

Principal Findings. Most physicians perceived the U.S. population CHD prevalence as higher in men (48.4 percent) or similar by gender (44.9 percent). For the observed patient, 52 percent did not change their CHD diagnostic certainty based on patient gender. Forty-eight percent of physicians were inconsistent in their population-level and individual-level CHD assessments. Physicians’ assessments of CHD prevalence did not attenuate the observed gender effect in diagnostic certainty for the individual patient.

Conclusions. Given an adequate presentation of CHD symptoms, physicians may deviate from their prevalence data during diagnostic decision making. Physicians’ priors on CHD prevalence did not explain the gender effect in CHD certainty. Future research should examine personal stereotypes as an explanation for gender differences.

Key Words. Clinical decision making, coronary heart disease, diagnosis, gender differences in care, statistical discrimination

Variations in clinical decision making have been evident for over 20 years, prompting efforts in research and training to minimize disparities in diagnostic and treatment decisions. While guidelines for disease management are thought to be useful in regulating patient care, the initial diagnostic decision making process inherently remains more difficult to standardize (McKinlay et al. 2006, 2007; Bonte et al. 2008). When interpreting a patient's symptoms, physicians undergo numerous complex cognitive processes to reach an active diagnosis (Hershey and Baron 1987; Ferreira et al. 2006; Krynski and Tenenbaum 2007). These processes involve filtering large amounts of potentially conflicting information and incorporating relevant evidence to come to a diagnostic decision. To help, commonly advised methods use epidemiologic prevalence data in probability analysis or evidence-based decision support tools (Diamond et al. 1980, 1983; Halkin et al. 1998; Reynolds 2001). These methods entail weighing the patient's symptoms along with the prior likelihood that the patient has a condition, given other background information. Two major funnels of information that remain key for clinical decision making during an initial patient encounter, whether or not formal decision aids are used, are (1) details on the individual patient's presentation and (2) prior knowledge on likelihood of a disease, based on prevalence data or clinical experience with groups similar to the patient.

Ideally, the amount of weight a physician places on each of these two funnels of information varies depending on the quality of the information available from each. In the situation where high-quality data are available from both sources, patient-specific clinical information should be considered more heavily than preexisting prevalence data (Lutfey et al. in press). Prevalence data indicate population risks, not specific individual risks; furthermore, epidemiologic data necessitate statistical assumptions and are, in essence, derived from clinical decision making (Lutfey et al. in press; Rockhill, Kawachi, and Colditz 2000). Consider as an illustration the extreme case where the patient's symptoms are clearly communicated and unambiguously point the physician to a diagnosis; here, preexisting data suggesting that the prevalence of the suspected disease is low in this patient population should not override the obvious diagnosis signaling from the individual's presentation.

Address correspondence to Nancy N. Maserejian, Sc.D., New England Research Institutes Inc., 9 Galen Street, Watertown, MA 02453; e-mail: nmaserejian@neriscience.com. Karen E. Lutfey, Ph.D., and John B. McKinlay, Ph.D., are also with the New England Research Institutes Inc., Watertown, MA.

In the more common situation where the patient's symptoms do not provide an obvious diagnostic decision path, physicians may have to put more weight onto their prior probability of disease, or "priors" (Medicine 2003). In doing so, physicians use presumably accurate group data to help make a decision in light of uncertainty about the individual. The use of prior data in the face of uncertainty about the individual patient has been termed "statistical discrimination" (Balsa, McGuire, and Meredith 2005; McGuire et al. 2008). One hypothesized mechanism by which statistical discrimination affects clinical decision making, termed the "prevalence" hypothesis, is that physicians use prior data on the prevalence of disease to help determine their certainty of a diagnosis (i.e., posterior probability of disease). For example, a physician may believe that the prevalence of coronary heart disease (CHD) varies by gender, so when faced with some diagnostic uncertainty with a patient, the physician will consider the patient's gender as a factor. Thus, if a male and a female patient each present with exactly the same ambiguous CHD symptoms, a physician who refers to priors on the prevalence of CHD may be more certain of CHD in the male patient, for whom the population prevalence is higher (Rosamond et al. 2008). Another possible mechanism has been termed the "miscommunication" hypothesis, in which statistical discrimination occurs because the physician has trouble understanding certain patients, perhaps because of language, culture, or communication patterns (Balsa, McGuire, and Meredith 2005). With miscommunication, the physician misses the patient's signals and must rely more heavily on prior data. For example, if women tend to use vague language when describing their CHD symptoms, physicians may be more likely to miss a relevant diagnosis with women. It should be noted that within either of these hypotheses lies the potential for some role of personal stereotypes or prejudices, which is generally regarded as unjustifiable during clinical decision making (Balsa, McGuire, and Meredith 2005).

Considering the complicated pathways to a diagnostic decision, recent experiments investigating disparities in clinical decision making have attempted to control for the miscommunication pathway (Arber et al. 2006; McKinlay et al. 2006; Bonte et al. 2008; Lutfey et al. 2008). In these experiments, physicians observe simulated patients who have the same exact verbal and nonverbal symptom presentation but vary in key factors of interest, such as gender, age, or race. In the case of CHD, results showed that even with the exact same patient communication of cardinal symptoms, physician certainty of the CHD diagnosis is significantly greater if the observed patient is male—even holding other factors (age, race, and socioeconomic status) constant (Arber et al. 2006; Bonte et al. 2008; Lutfey et al. 2008). It is plausible that

statistical discrimination, via the prevalence hypothesis, may explain all or part of this gender effect.

Understanding the cognitive processes of diagnostic decisions and distinguishing which, if any, types of discrimination are most involved, may be critical to understand health disparities and suggest possible interventions (Lutfey and Ketcham 2005). To help understand how prevalence data and statistical discrimination may affect clinical decision making for CHD, our specific objectives were as follows:

1. Examine physicians' perceptions of the prevalence of CHD in men versus women and assess whether their posterior probability of CHD for individual patients (as measured by their diagnostic certainty of CHD) is consistent with these perceptions.
2. Test the hypothesis that physician or patient characteristics are associated with consistency (as assessed in Objective 1).
3. Test the hypothesis that statistical discrimination, via the prevalence hypothesis, can explain the previously reported gender effect in clinical decision making for CHD.

CHD is the single greatest cause of death for men and women in the United States and Europe. The case of CHD is particularly well-suited for our purposes because epidemiologic prevalence data on CHD are extensive and can be regarded as a reliable and sufficiently specific source of priors; also, it generally shows that the prevalence of CHD is higher in men than in women (Rosamond et al. 2008). As our focus is on understanding the role of priors, we use data from an experiment that restricted the role of the individual patient's symptoms and was designed to eliminate differential miscommunication by standardizing all patients' presentations to be strongly suggestive of CHD. Objective 1 descriptively examines the role of prevalence data in clinical decision making among these physicians as an initial step. It attempts to inform the question, do physicians hold on to their priors when faced with an individual patient presentation strongly suggestive of CHD? Objective 2 investigates which physicians are most likely to maintain priors, and with which patients, to help inform targets for interventions in clinical decision making. Investigating physicians' adherence to their prior probabilities during clinical decision making is important to decipher cognitive processes; furthermore, it may reveal the extent to which gender differences in CHD are due to reliance on prior data.

METHODS

Data Source

Our data source was the NIH-funded project “Cognitive Basis of CHD Disparities.” This project conducted a factorial experiment to measure the unconfounded effects of (a) patient attributes (age, gender, race, and socioeconomic status); (b) physician characteristics (gender and years of clinical experience); and (c) cognitive priming status on medical decision making for an actor “patient” presenting with CHD in a videotaped vignette. The study was approved by New England Research Institutes’ Institutional Review Board and obtained signed informed consent from all participants.

Participants

Physicians were sampled from North and South Carolina to fill four design strata (gender by experience) totaling 256 participants. Eligibility criteria were as follows: (a) internist, general practitioner, or family practitioner; (b) graduated from medical school between 1996 and 2001 or between 1960 and 1987 (to preserve orthogonality and ensure clear separation on level of clinical experience); and (c) currently seeing patients at least half-time. Prospective participants were sent letters and screened for eligibility by telephone. An appointment was scheduled with each eligible participant for a one-on-one, structured interview. Interviews were conducted over a 10-month period in 2006–2007.

Clinical Presentation: “Patient” Video Vignette

Professional actors portrayed patients presenting to a primary care provider with the essential signs and symptoms of CHD. Sixteen versions of the scenario were created using eight actors/actresses, systematically varying the patient’s gender, age (55 or 75 years), race (black or white), and socioeconomic status (lower or upper, as depicted by occupation). The script was designed to include the key diagnostic evidence that would lead physicians to suspect CHD, as described in Table 1.

To investigate the effect of cognitive priming status, half of physicians were randomly selected to be primed (i.e., explicitly directed) to consider a CHD diagnosis before viewing the vignette. These participants were told that a physician who had seen the patient while s/he was on vacation had mentioned the possibility of CHD and suggested s/he see her/his primary care physician upon returning home.

Table 1: Symptoms Embedded in the Clinical Scenario

Classic symptoms and signs of coronary heart disease

- Chest pain, described as:
 - worsening with exertion
 - after eating a big or spicy meal
 - when in a stressful situation
 - relieved by rest
 - feels different from heartburn
 - duration of about 3 months
- Pain in the back between the shoulder blades
- High blood pressure
- “Levine fist”—clenched fist to the sternum, as a nonverbal indicator of cardiac pain
- Stress

*Additional symptoms and patient complaints**

- Gastrointestinal discomfort (e.g., feeling “full all the time,” “gassy”)
- Heartburn
- Indigestion not relieved with antacids
- Spouse says patient is not acting like her/himself lately
- Mood changes: easily irritated, concerned
- Low energy level

*Since patients seldom present as clear-cut textbook cases, additional symptoms and complaints that are not exclusively indicative of CHD were also presented. The purpose was not to make the physicians’ diagnostic task more difficult, but to increase the clinical authenticity of the scenario, so that it more accurately represented how actual patients present (Milner et al. 1999; DeVon et al. 2008; Dey et al. 2008).

Measurement of CHD Certainty for a Male versus Female Patient

After viewing the vignette, a trained interviewer asked the physician to say what she or he thought was going on with the patient, and to state a diagnostic certainty value (on a scale of 0–100, 0 indicating no certainty and 100 indicating complete certainty) for each diagnosis. The diagnostic certainty of CHD indicates a posterior probability. The majority of physicians (98.8 percent) identified CHD as one possible diagnosis; physicians who did not mention CHD were asked later in the interview to state their CHD diagnostic certainty. Then, in a gender substitution exercise, physicians were asked what their CHD certainty value would be if a patient of the opposite gender had presented in the exact same way as their observed patient.

Measurement of Prior Data

Physicians were asked to provide their estimate of the prevalence of CHD in the entire U.S. adult population. They were then asked to rate the accuracy of a prevalence estimate of 6.9 percent for adult men and women as being “far too

low, [men/women] have a much higher base rate,” or “far too high, [men/women] have a much lower base rate” as the extremes of a seven-point Likert scale, without being told that this was the overall prevalence of CHD published by the American Heart Association in 2006 (Thom et al. 2006).

Analytic Strategy

In accordance with our objectives, the analysis was conducted in three phases. First, we calculated descriptive statistics for physician characteristics and responses regarding prior data and diagnostic certainty. To investigate the cognitive decision making process, a primary interest in this analysis is whether physicians were consistent in the diagnostic process regarding their CHD priors and diagnostic certainties. We categorized physicians as “consistent” or “inconsistent” in the diagnostic process based on the relationship between the direction of change in CHD prevalence estimates and the direction of change in CHD diagnostic certainty for the individual patient when asked to substitute patient gender. That is, a physician was labeled inconsistent if his/her group (i.e., prevalence) and individual (i.e., patient) assessments were not concordant (e.g., if a physician perceived the prevalence as higher in men, but was equally or more certain of the CHD diagnosis for the individual patient if she were female, then the physician would be labeled inconsistent). For our purposes, the measure of inconsistency is not meant to indicate proper or improper decision making; rather, it serves to reveal a cognitive aspect of the decision making process.

For our second objective, we examined predictors of consistency using statistical models. In models using only design factors (patient factors of gender, age, race, socioeconomic status; physician factors of gender, experience, priming status), the balanced factorial design allows the unconfounded estimation of main effects and interactions using analysis of variance. In models additionally evaluating nondesign physician factors, such as attitudes regarding medical literature and published prevalence rates, results were similar in unadjusted and multivariate models.

Finally, we examined whether a previously reported effect of patient gender on CHD certainty (Arber et al. 2006; Bonte et al. 2008; Lutfey et al. 2008) may be explained by physicians’ priors (i.e., the prevalence hypothesis of statistical discrimination). Here, we compared the effect of gender in two generalized linear models shown below: (1) the basic factorial model, in which the dichotomous patient and physician characteristics are independent variables, and certainty of CHD is the dependent continuous variable (Y_i); and

(2) Model 1 with added categorical indicator variables for the physicians' perceptions of CHD population prevalence (categories of higher in men, higher in women, or, the reference group, similar for men and women).

$$Y_i = \alpha + \beta_{pg} \text{PatientGender}_i + \beta_{pa} \text{PatientAge}_i + \beta_{pr} \text{PatientRace}_i + \beta_{ses} \text{PatientSES}_i + \beta_{pgs} \text{PhysicianGender}_i + \beta_{pe} \text{PhysicianExperience}_i + \beta_{ps} \text{PrimingStatus}_i + e_i \quad (1)$$

$$Y_i = \alpha + \beta_{hm} \text{HigherMen}_i + \beta_{hw} \text{HigherWomen}_i + \beta_{pg} \text{PatientGender}_i + \beta_{pa} \text{PatientAge}_i + \beta_{pr} \text{PatientRace}_i + \beta_{ses} \text{PatientSES}_i + \beta_{pgs} \text{PhysicianGender}_i + \beta_{pe} \text{PhysicianExperience}_i + \beta_{ps} \text{PrimingStatus}_i + e_i \quad (2)$$

We include a pure error term with 128 degrees of freedom because there were two replications of the experiment. There is evidence for statistical discrimination if the effect of patient gender as observed in model 1 is reduced or eliminated in model 2. In addition, in exploratory analysis stratified by patient gender, we included physician estimates of the overall U.S. population prevalence of CHD as a continuous variable, to explore whether the role of this prior data differed for male and female patients.

Because priming status has high potential to affect physicians' cognitive processes during decision making, we repeated all analyses separately for primed and unprimed physicians. Results were similar to the main analysis, which adjusted for this design factor in all multivariate models, and so are not reported.

All statistical tests were two sided and performed at an α level of 0.05 using SAS v. 9.1 (SAS Institute, Cary, NC).

RESULTS

Objective 1: Perceptions of CHD Population Prevalence, Diagnostic Certainty, and Inconsistency

Physicians most commonly perceived the overall U.S. adult population prevalence of CHD to be higher in men (48.1 percent), though a substantial 44.9 percent thought the prevalence was similar by gender (Table 2). The majority of physicians provided an estimate of the overall CHD prevalence that was higher than the rate published by the American Heart Association (data not shown) (Thom et al. 2006).

For their observed patient, physicians were significantly more certain of a CHD diagnosis if the patient were male (mean certainty on scale of

Table 2: Descriptive Statistics for Physicians' Responses to CHD Population Prevalence Assessments and Diagnostic Certainty for the Individual Patient

Physicians' assessments of prevalence of CHD in U.S. adult population (%)	
Higher in men	48.1
Higher in women	7.0
Similar for men and women	44.9
Certainty of CHD for the observed patient (mean, on a scale of 0–100)	
Female patient	53.0
Male patient	61.7
Physicians' change in certainty of CHD for the observed patient, with patient gender substitution (%)	
Higher in a male patient	32.4
Higher in a female patient	16.0
Similar for a male or female patient	51.6
Consistency between physician estimate of prevalence and CHD certainty for the individual patient regarding gender effects (%)	
Physicians who were consistent	51.6
Physicians who were inconsistent	48.4

CHD, coronary heart disease.

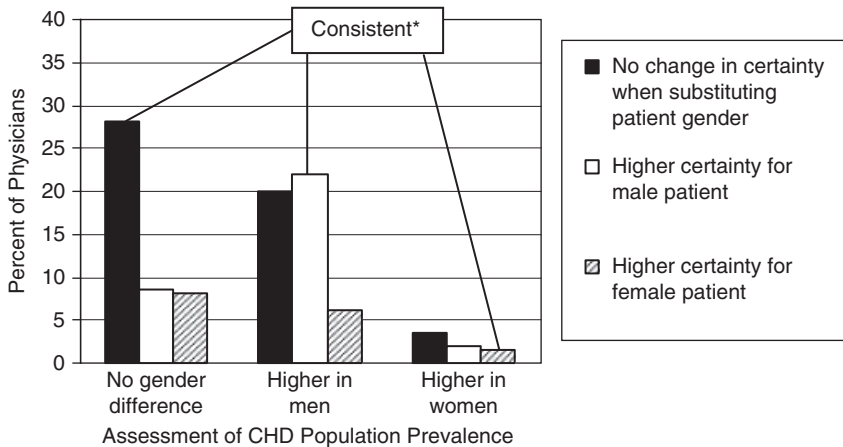
0–100: 61.7 versus 53.0, $p = .002$). When asked how their certainty of CHD would change if the patient had been male versus female, 52 percent of all physicians said their certainty would not change, 32 percent would be more certain for male patients, and 16 percent more certain for female patients.

Inconsistency between perceptions of prevalence for men versus women and diagnostic certainty for the individual patient occurred with almost half (48.4 percent) of physicians. Inconsistent physicians most frequently diverged from their priors to diagnose male and female patients with equal certainty in the vignette (Figure 1). Physicians who thought that the prevalence of CHD was either higher in men or higher in women were more likely to be inconsistent than those who thought that there was no gender difference ($p = .001$, Table 3). Although there was considerable variation in the responses, overall, the most common response group was physicians who were consistent in assessing no gender difference in population prevalences and no gender difference in their certainty of CHD for the given patient (28 percent of all participants, Figure 1).

Objective 2: Physician or Patient Characteristics as Predictors of Consistency

In unadjusted and multivariate models, no physician characteristics were significantly associated with being consistent in the assessment of CHD population prevalence and diagnostic certainty for the given patient (Table 3).

Figure 1: Change in Certainty of CHD Diagnosis for a Male versus Female Individual Patient, by Physician Assessment of CHD Prevalence in the U.S. Population



*Consistent between the assessment of CHD population prevalence and certainty of the CHD diagnosis for the observed individual male or female patient.

Similarly, patient factors of age, race, gender, or socioeconomic status were not associated with consistency (data not shown).

Objective 3: Statistical Discrimination as an Explanation for Gender Effects in CHD Certainty

As this experiment and others (Arber et al. 2006; Bonte et al. 2008; Lutfey et al. 2008) have found significant effects of patient gender on CHD certainty, we analyzed whether adjustment for physicians' assessments of gender-based priors explains or attenuates the role of the observed patient's gender. Multivariate model results showed that physicians' gender-based population prevalence assessments were not associated with CHD certainty ($p = .5$), and the effect of patient gender remained strong (adjusted mean CHD certainty: 63.6 for male versus 54.7 for female patients, $p = .001$). Thus, this analysis provided no evidence to support the prevalence hypothesis as an explanation for statistical discrimination by gender.

Interestingly, in exploratory analyses of physicians' absolute estimates of the overall adult CHD prevalence, there was a statistically significant asso-

Table 3: Percent of Physicians Consistent in Their Gender Assessments of CHD Population Prevalence and Their Diagnostic Certainty for the Individual Patient, by Physician Characteristics*

<i>Physician Characteristics</i>	<i>n</i>	<i>% Consistent</i>	<i>p Value</i>
Gender			
Male	128	53.1	.62
Female	128	50.0	
Clinical experience			
More	128	48.4	.32
Less	128	54.7	
Primed to consider CHD as a diagnosis			
Yes	128	49.2	.45
No	128	53.9	
Reported following medical literature closely			
Yes	61	57.4	.30
No	195	49.7	
Believes published prevalences are accurate			
Yes	117	53.0	.68
No	139	47.0	
Believes his/her patient population is similar to the general U.S. adult population			
Yes	93	52.7	.79
No	163	47.3	
Assessment of CHD prevalence in general U.S. adult population			
No gender difference	115	62.6	.001
Higher in men	123	45.5	
Higher in women	18	22.2	

*Unadjusted χ^2 *p* values; adjustment for other physician or patient factors did not appreciably affect results.

CHD, coronary heart disease.

ciation between this population estimate and diagnostic certainty for the individual patient only for male patients ($p = .02$; female patients $p = .52$). Physicians who assigned a higher overall population prevalence were more likely to ascribe a higher certainty to the CHD diagnosis for the observed patient only if the patient were male, suggesting that these prior data were not as relevant in the diagnostic decision for female patients.

DISCUSSION

In this series of analyses, we have shown that physicians often deviated from their preconceived notions of the likelihood of disease when diagnosing

patients, thereby placing more weight on the current patient's presentation and less weight on prior probabilities. Patients in our experiment presented with cardinal symptoms of CHD, such that the level of evidence provided by the symptoms led the vast majority of physicians to consider CHD as a diagnosis, and just over half of the physicians (51.6 percent) reported that they would not change their diagnostic certainty based on patient gender. However, in the main factorial experiment, physicians were significantly less certain of the CHD diagnosis for female patients. Our finding that this gender effect could not be explained by the physicians' prior notions of CHD probabilities indicates that statistical discrimination via the prevalence hypothesis was not the underlying reason for gender differences in CHD certainty.

Why women were diagnosed with a lower certainty, despite presenting with the exact same symptoms and controlling for gender-relative priors in CHD prevalence, is critical to understand if worrisome inequalities in clinical decision making and health care are to be appropriately addressed. One possibility is that physicians behaved differently for men and women because of personally held stereotypes or prejudices. Discrimination resulting from personal stereotypes is very different from statistical discrimination resulting from the application of prior probabilities (Balsa, McGuire, and Meredith 2005; McGuire et al. 2008). That is, when physicians use prior probabilities to guide decisions, they are attempting to use as much information as they have available to guide their decisions, in the best interests of the patients. When they are influenced by personal prejudices, they are not acting in the best interests of their patients (Balsa, McGuire, and Meredith 2005). In the current study, the extent to which personal stereotypes or prejudices may explain our results is unknown.

Another possible reason that we did not see evidence of statistical discrimination underlying the gender effect in CHD diagnosis certainty is that our measure of prior information may not have sufficiently captured the priors that physicians held. First, it would have been helpful to know the physicians' clinical experiences with male versus female patients and which CHD symptoms they typically encountered. It is possible that physicians with more frequent exposure to male patients with CHD symptoms similar to the vignette would be more certain of the male simulated patient's diagnosis. Second, our analysis is based on a relative comparison of the physicians' prevalence estimates for men and women, rather than absolute values of estimates for each. However, a relative comparison may be preferable, particularly considering that the physicians' absolute estimates of the overall population prevalence of CHD were substantially higher than the published rate of 6.9 percent (Thom et al. 2006). Assuming that published prevalence data are correct, other studies

have also found that prior probabilities estimated by physicians were inaccurate, to the extent that the authors suggested that the use of prior probabilities as a tool for clinical decision making might cause more harm than benefit (Cahan et al. 2003). Important to note is that in our study, the overall estimate was associated with the CHD diagnostic certainty for male patients only; among female patients, the role of this prior information was irrelevant. This difference suggests that the symptom presentation held more weight for female patients than it did for male patients. Alternatively, physicians may have been more confident of the relevance or accuracy of CHD population rates for male patients. Because the CHD prevalence estimate was important in the diagnosis of male patients, there may have been stronger overall evidence—both priors and current patient presentation—to increase the certainty of the CHD diagnosis in male patients.

While information on absolute rates of disease may be helpful for clinical decision making, understanding disparities in health care requires examining relative differences across sociodemographic groups. Comparing genders, the majority of physicians in our study thought that the prevalence estimate was similar for men and women, although a sizeable 46 percent believed it was higher in men. Our finding that the minority of physicians who assessed a higher CHD prevalence for women were most likely to be inconsistent with this notion in their diagnostic certainty suggests that the symptom presentation strongly outweighed their prior beliefs. Accordingly, physicians whose priors held that CHD prevalence was similar by gender were most likely to be consistent. Physician level of clinical experience, keeping up with medical literature, beliefs in the accuracy of published prevalence rates, and priming status did not help predict which physicians would be more likely to adhere to their priors in the diagnostic process.

For present purposes, a critical benefit of using the experimental vignette is that it allows for the manipulation of several variables at once, thereby providing unconfounded results for factors (e.g., race and socioeconomic status) that are otherwise nearly impossible to disentangle (i.e., ensuring internal validity). Studies comparing the vignette methodology with standardized patients and other methods have shown that vignettes are also externally valid for studies of medical decision making and assessments of quality of care (Braspenning and Sergeant 1994; Peabody et al. 2000; Veloski et al. 2005; Robra et al. 2006). To further enhance the external validity of our results (i.e., that physicians behave similarly under experimental conditions as in everyday clinical practice), we took three precautionary steps. First, considerable effort was devoted to ensure the clinical authenticity of the videotaped pre-

sentation. This was achieved by basing the scripts on clinical experience of physician advisors, filming with experienced clinicians present, and by using highly trained professional actors/actresses. Second, physicians viewed the vignette in the context of their practice day (not at a professional meeting, course update, or home) so that it was likely they encountered real patients before and after they viewed the patient in the videotape, thereby retaining as much of the situational context as possible. Third, physicians were specifically instructed at the outset to view the patient as one of their own and to respond as they would typically respond in their own practice. When asked if the patient viewed on the videotape was typical of patients they encounter in everyday practice, 90 percent considered them very typical or reasonably typical.

In this paper, we have examined two major funnels of information that may help a physician come to a diagnostic decision. A critical issue in clinical decision making is that both funnels of information are subject to some unknown level of error. For example, reports of a given patient's symptoms may be faulty if there is miscommunication between the doctor and the patient, or if the patient does not provide certain details, perhaps because of embarrassment or a belief that such details are irrelevant. An advantage of our experiment was that it minimized miscommunication to the greatest extent possible; while it is impossible to control physicians' perceptions of the patient's signals, all facets of the patient presentation (apart from design factors such as gender) were exactly the same across all patient encounters.

The second funnel, which includes published data in the medical literature or personal physician knowledge or experience, may suffer its own biases. For instance, data on base rates ultimately stem from reports of physicians' diagnoses, which are prone to error. Thus, for use of prevalence data to be an acceptable option, it is essential that data are accurate and up to date. For example, if epidemiological data indicating that the prevalence of CHD is higher in men were erroneous, then using such data to dismiss a CHD diagnosis in a female patient in the face of symptom uncertainty would lead to a faulty diagnosis and delayed treatment.

We have shown that given a sufficient symptom presentation, prevalence data were often outweighed during the decision making process. Thus, our results suggest that we need not worry excessively over the potential that statistical discrimination needlessly affects published rates for CHD by gender, given an adequate patient symptom assessment. Reliance on prior beliefs rather than patient-specific information to guide clinical decision making is difficult to justify when high-quality patient information and low-cost tests are available (Balsa, McGuire, and Meredith 2005). The finding that physicians'

perceptions of base rates may be inaccurate further supports this notion. While our methodology cannot definitively rule out the potential for any role of priors in the gender effect in CHD diagnoses, the influence of stereotypes or prejudice should be examined in future work. In addition, the extent to which physicians rely on priors to diagnose patients when patient presentation is less informative or when miscommunication is likely is critical to examine in the field of clinical decision making.

ACKNOWLEDGMENTS

Joint Acknowledgment/Disclosure Statement: This study was supported by National Heart, Lung, and Blood Institute grant R01 HL079174. The authors wish to thank Dr. Carol Link, senior statistician, for data preparation and statistical guidance.

Disclosures: There are no conflicts of interest or disclosures.

Disclaimers: There are no disclaimers.

REFERENCES

- Arber, S., J. McKinlay, A. Adams, L. Marceau, C. Link, and A. O'Donnell. 2006. "Patient Characteristics and Inequalities in Doctors' Diagnostic and Management Strategies Relating to CHD: A Video-Simulation Experiment." *Social Science Medicine* 62 (1): 103–15.
- Balsa, A. I., T. G. McGuire, and L. S. Meredith. 2005. "Testing for Statistical Discrimination in Health Care." *Health Services Research* 40 (1): 227–52.
- Bonte, M., O. von dem Knesebeck, J. Siegrist, L. Marceau, C. Link, S. Arber, A. Adams, and J. B. McKinlay. 2008. "Women and Men with Coronary Heart Disease in Three Countries: Are They Treated Differently?" *Womens Health Issues* 18 (3): 191–8.
- Braspenning, J., and J. Sergeant. 1994. "General Practitioners' Decision Making for Mental Health Problems: Outcomes and Ecological Validity." *Journal of Clinical Epidemiology* 47 (12): 1365–72.
- Cahan, A., D. Gilon, O. Manor, and O. Paltiel. 2003. "Probabilistic Reasoning and Clinical Decision-Making: Do Doctors Overestimate Diagnostic Probabilities?" *Quarterly Journal of Medicine* 96 (10): 763–9.
- DeVon, H. A., C. J. Ryan, A. L. Ochs, and M. Shapiro. 2008. "Symptoms across the Continuum of Acute Coronary Syndromes: Differences between Women and Men." *American Journal of Critical Care* 17 (1): 14–24; quiz 25.
- Dey, S., M. D. Flather, G. P. Devlin, D. Brieger, E. P. Gurfinkel, P. G. Steg, G. Fitzgerald, E. A. Jackson, and K. A. Eagle. 2008. "Sex-Related Differences in the Presentation, Treatment and Outcomes among Patients with Acute Coro-

- nary Syndromes: The Global Registry of Acute Coronary Events." *Heart* 95 (1): 20–6.
- Diamond, G. A., J. S. Forrester, M. Hirsch, H. M. Staniloff, R. Vas, D. S. Berman, and H. J. Swan. 1980. "Application of Conditional Probability Analysis to the Clinical Diagnosis of Coronary Artery Disease." *Journal of Clinical Investigation* 65 (5): 1210–21.
- Diamond, G. A., H. M. Staniloff, J. S. Forrester, B. H. Pollock, and H. J. Swan. 1983. "Computer-Assisted Diagnosis in the Noninvasive Evaluation of Patients with Suspected Coronary Artery Disease." *Journal of American College of Cardiology* 1 (2 Pt 1): 444–55.
- Ferreira, M. B., L. Garcia-Marques, S. J. Sherman, and J. W. Sherman. 2006. "Automatic and Controlled Components of Judgment and Decision Making." *Journal of Personality and Social Psychology* 91 (5): 797–813.
- Halkin, A., J. Reichman, M. Schwaber, O. Paltiel, and M. Brezis. 1998. "Likelihood Ratios: Getting Diagnostic Testing into Perspective." *Quarterly Journal of Medicine* 91 (4): 247–58.
- Hershey, J. C., and J. Baron. 1987. "Clinical Reasoning and Cognitive Processes." *Medical Decision Making* 7 (4): 203–11.
- Institute of Medicine. 2003. *Unequal Treatment: Confronting Racial and Ethnic Disparities in Health Care*. Washington, DC: National Academies Press.
- Krynski, T. R., and J. B. Tenenbaum. 2007. "The Role of Causality in Judgment under Uncertainty." *Journal of Experimental Psychology: General* 136 (3): 430–50.
- Lutfey, K., C. Link, L. Marceau, and J. McKinlay. In press. "Sociological Contributions to the Study of Clinical Decision Making." In *Encyclopedia of Medical Decision Making*, edited by M. Kattan and M. Cowan. New York: Russell Sage.
- Lutfey, K. E., and J. D. Ketcham. 2005. "Patient and Provider Assessments of Adherence and the Sources of Disparities: Evidence from Diabetes Care." *Health Services Research* 40 (6 Pt 1): 1803–17.
- Lutfey, K. E., C. L. Link, R. W. Grant, L. D. Marceau, and J. B. McKinlay. 2008. "Is Certainty More Important Than Diagnosis for Understanding Race and Gender Disparities?: An Experiment Using Coronary Heart Disease and Depression Case Vignettes." *Health Policy* 89 (3): 279–87.
- McGuire, T. G., J. Z. Ayanian, D. E. Ford, R. E. Henke, K. M. Rost, and A. M. Zaslavsky. 2008. "Testing for Statistical Discrimination by Race/Ethnicity in Panel Data for Depression Treatment in Primary Care." *Health Services Research* 43 (2): 531–51.
- McKinlay, J., C. Link, L. Marceau, A. O'Donnell, S. Arber, A. Adams, and K. Lutfey. 2006. "How Do Doctors in Different Countries Manage the Same Patient? Results of a Factorial Experiment." *Health Services Research* 41 (6): 2182–200.
- McKinlay, J. B., C. L. Link, K. M. Freund, L. D. Marceau, A. B. O'Donnell, and K. L. Lutfey. 2007. "Sources of Variation in Physician Adherence with Clinical Guidelines: Results from a Factorial Experiment." *Journal of General Internal Medicine* 22 (3): 289–96.

- Milner, K. A., M. Funk, S. Richards, R. M. Wilmes, V. Vaccarino, and H. M. Krumholz. 1999. "Gender Differences in Symptom Presentation Associated with Coronary Heart Disease." *American Journal of Cardiology* 84 (4): 396–9.
- Peabody, J. W., J. Luck, P. Glassman, T. R. Dresselhaus, and M. Lee. 2000. "Comparison of Vignettes, Standardized Patients, and Chart Abstraction: A Prospective Validation Study of 3 Methods for Measuring Quality." *Journal of the American Medical Association* 283 (13): 1715–22.
- Reynolds, T. 2001. "Disease Prediction Models Aim to Guide Medical Decision Making." *Annals in Internal Medicine* 135 (8 Pt 1): 637–40.
- Robra, B. P., H. Kania, O. Kuss, K. Schonfisch, and E. Swart. 2006. "Determinants of Hospital Admission—Investigation by Case Vignettes." *Gesundheitswesen* 68 (1): 32–40.
- Rockhill, B., I. Kawachi, and G. A. Colditz. 2000. "Individual Risk Prediction and Population-Wide Disease Prevention." *Epidemiological Review* 22 (1): 176–80.
- Rosamond, W., K. Flegal, K. Furie, A. Go, K. Greenlund, N. Haase, S. M. Hailpern, M. Ho, V. Howard, B. Kissela, S. Kittner, D. Lloyd-Jones, M. McDermott, J. Meigs, C. Moy, G. Nichol, C. O'Donnell, V. Roger, P. Sorlie, J. Steinberger, T. Thom, M. Wilson, and Y. Hong. 2008. "Heart Disease and Stroke Statistics–2008 Update: A Report from the American Heart Association Statistics Committee and Stroke Statistics Subcommittee." *Circulation* 117 (4): e25–146.
- Thom, T., N. Haase, W. Rosamond, V. J. Howard, J. Rumsfeld, T. Manolio, Z. J. Zheng, K. Flegal, C. O'Donnell, S. Kittner, D. Lloyd-Jones, D. C. Goff Jr., Y. Hong, R. Adams, G. Friday, K. Furie, P. Gorelick, B. Kissela, J. Marler, J. Meigs, V. Roger, S. Sidney, P. Sorlie, J. Steinberger, S. Wasserthiel-Smoller, M. Wilson, and P. Wolf. 2006. "Heart Disease and Stroke Statistics–2006 Update: A Report from the American Heart Association Statistics Committee and Stroke Statistics Subcommittee." *Circulation* 113 (6): e85–151.
- Veloski, J., S. Tai, A. S. Evans, and D. B. Nash. 2005. "Clinical Vignette-Based Surveys: A Tool for Assessing Physician Practice Variation." *American Journal of Medical Quality* 20 (3): 151–7.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of this article:

Appendix SA1: Author Matrix.

Please note: Wiley-Blackwell is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.