
What happens along the diagnostic pathway to CHD treatment? Qualitative results concerning cognitive processes

Karen E. Lutfey and John B. McKinlay

New England Research Institutes, Watertown MA, United States

Abstract Extensive research on health disparities documents persistent differential diagnosis and treatment of many conditions according to patient characteristics, physician attributes, and healthcare systems. Less is known about *how* physicians arrive at their decisions. We use qualitative data from a vignette-based factorial experiment to examine how physicians reason through and account for their clinical decisions, and how variations arise despite the presentation of identical symptoms of coronary heart disease (CHD). We find that physicians show evidence of cognitive biases but also actively interpret social characteristics they deem relevant to medical treatment. In an uncertain clinical context, these diagnostic pathways expose key junctures wherein physicians are detoured to alternative diagnoses, their certainty of CHD lowered, and scientific logic makes it difficult to return to a CHD diagnosis — thereby providing a fuller picture of why some cases are counted as CHD while others are not. These results have important implications insofar as diagnostic decisions like these contribute to the compilation of epidemiologic base rates, and are therefore used as part of Bayesian decision making to determine the probability of CHD in subsequent patients. This work resonates with social constructivist concerns regarding the ways disease categories are established and maintained, and potential sources of bias in official rates detected.

Keywords: clinical decision making, qualitative methods, coronary heart disease, social constructionism

Over several decades, medical, sociological, and epidemiological studies have documented disparities in the prevalence, treatment, and outcomes of coronary heart disease (CHD), both in the US and cross-nationally (Gurjeva *et al.* 2005, Weisz, Gusmano and Rodwin 2004). Research on differential diagnostic and treatment decisions for CHD patients has shown differences in CHD management by patient characteristics (Armstrong, Strogatz and Wang 2004, Popescu, Vaughan-Sarrazin and Rosenthal 2007), physician attributes (Britt *et al.* 1996, McKinlay *et al.* 2002), healthcare settings (Bird *et al.* 2007, Curoe, Kralewski and Kaissi 2003), and geographic location (Goodman and Grumbach 2008, von dem Knesebeck *et al.* 2008).

As one facet of health disparities, variation in clinical decision making (CDM) affecting patient diagnosis and treatment has been well documented across many conditions, ranging

from breast cancer (Krupat *et al.* 1999, Maly *et al.* 2006) to heart disease (Adams *et al.* 2006, Arber *et al.* 2004) to depression (Anglin, Link and Phelan 2006, McKinlay *et al.* 2002). Furthermore, this variability is also manifest in different aspects of clinical decision making, with differences observed not only in diagnosis (Freund *et al.* 2003, Schulman *et al.* 1999), but also treatment (Barnhart and Wassertheil-Smoller 2006, Maly *et al.* 2006), referrals (Forrest *et al.* 2006, Starfield *et al.* 2002), physical examinations, test ordering, and medical interviewing (McKinlay *et al.* 2007).

Despite the extensive documentation of variations in diagnosis and treatment, less is known about the mental and social reasoning processes that lead physicians to make these disparate decisions. Psychologists have documented at least 30 different types of subconscious cognitive bias that may affect clinical reasoning (Croskerry 2003, Galanter and Patel 2005), and sociological work has shown that diagnostic certainty — as distinct from identifying a specific diagnosis — is a significant predictor of clinical actions (Lutfey *et al.* 2009). Furthermore, a long tradition of sociological research has documented how uncertainty and mistakes are an indigenous feature of the work process as it unfolds (Paget 1988), how providers are socialised to manage such uncertainty, and the role of professional collegiality and internal social control in managing risks associated with it (Bosk 1979, Fox 1957). Questions remain, however, about the natural history of clinical decision making — that is, *how* physicians reason through cases and account for their decisions. Existing research does not fully capture how types of bias interact and comingle in the same case; the range of patient cues physicians identify as medically relevant and why; how patients' social characteristics are interpreted as relevant for clinical actions; and how uncertainty guides these processes in ways that lead to the disparities widely observed in existing literature. Examination of these diagnostic pathways is critical for learning why some patients are not diagnosed (or treated) with CHD despite presenting with identical symptoms. As Freidson remarks in *Profession of Medicine*:

Disease is put into a special category by sociologists because they mistake extensive social consensus for facts independent of consensus; yet a fact is established only by human consensus...But that the consensus is high does not make it any the less a social construction. In the case of physical illness, consensus is so extensive and taken for granted that we are inclined to impute it to a reality independent of our agreement (1970: 214–15).

Our analysis underscores the import of social processes in generating differential diagnostic and treatment outcomes in a context where symptom presentation has been standardised. The variations observed extend our knowledge to show how even the 'special category' of CHD does not have a consistent reality independent of its social environment and interpretation.

Methods

The qualitative data presented in this analysis are part of a larger video vignette factorial experiment study of physician decision making. The primary manipulation in the experiment was to prime half the physicians to consider a CHD diagnosis in order to determine whether underdiagnosis of some types of patients was due to physicians not considering CHD or purposely eliminating it from their differential diagnosis [unpublished manuscript]. The qualitative data analysed presently were collected for the purpose of analysing the cog-

native processes underlying physicians' decisions regarding the vignette they viewed in the experiment. The video vignette experiment simultaneously measured the 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 coronary heart disease. A full factorial of $2^4 = 16$ combinations of patient age (55 vs. 75), gender, race (Black vs. White) and socioeconomic status (SES) (lower vs. higher, depicted by current or former employment as a janitor or school teacher) were used for the video scenarios. One of the 16 combinations was shown to each physician. A total of 256 interviews (16 vignettes x four combinations of physician attributes x two priming conditions x two replications) were conducted over a period of 10 months in 2006–07. IRB approval was obtained and all participants signed informed consents.

To be eligible for selection, physicians had to: (a) be internists¹ or family practitioners with MD degrees (international medical graduates were included); (b) have graduated from medical school between either 1996–2001 or 1960–87; and (c) be currently providing primary care to patients in North or South Carolina more than half-time. The study was limited to generalist practitioners because, in the United States, these are the physicians who routinely care for the type of patient depicted in the vignette. While specialist physicians may exhibit different decision making patterns, they would be more likely to see such patients after preliminary diagnostic and testing work happens with a primary care provider. Physician graduation dates were chosen so that all physicians had had sufficient amounts of time since graduation to be acclimated to their work, but with the two groups far enough apart that there was clear separation between higher and lower levels of experience. Physicians were recruited to evenly fill cells representing the four types of physician attributes in the study, with even numbers of males and females with higher and lower levels of experience (64 in each group). Physicians were blind to all study hypotheses, including which condition was depicted in the vignette. They were told that it was a study of medical decision making and each participant was provided a stipend (\$200) to partially offset lost revenue and to acknowledge their participation.

A letter of introduction was mailed to prospective participants and screening telephone calls were conducted by in-house staff (in Massachusetts) to identify eligible physicians. For eligible physicians, an appointment was scheduled for a one-hour face-to-face interview, which was then conducted in the physician's office by one of two trained full-time field interviewers (one each in North and South Carolina). After each interview, field interviewers mailed all materials back to the in-house staff and were not involved in data analyses. To ensure quality control, interviewers were carefully trained and certified; selected interviews were tape recorded and reviewed by supervisors on a regular basis; monthly quality control conference calls were held with field interviewers, in-house training staff, and the Co-PI; and a Co-Investigator made a field visit to observe interviews in person.

Patients in the vignette presented with signs and symptoms suggestive of CHD, including chest pain of three months' duration that worsens with exertion or stress but is relieved with rest. There is also pain in the back between the shoulder blades and a report of elevated blood pressure at a previous visit to the doctor. To maximise clinical authenticity, the vignette script was developed from several tape-recorded role-playing sessions with experienced clinical advisors; these symptoms are highly consistent with those measured by the Rose Angina Questionnaire, a highly regarded and widely used survey measure of CHD (Bittner 2008). Based on their advice that live patients do not typically present as clear-cut textbook cases of specific conditions, the vignette also built in several distractions ('red herring' symptoms) potentially indicative of a gastrointestinal diagnosis. To this end,

the patient also complained of indigestion, feeling worse after a large or spicy meal, having pain similar to heartburn but unresponsive to antacids, and feeling full and 'gassy'. This was done not to make the physicians' diagnostic task more difficult, but to represent most accurately how actual patients present. The vignette also incorporates references to the patient's mood, including the spouse's report that the patient has been difficult to be around, the patient's self-report of feeling irritated and having decreased energy. The physician interacting with the patient in the vignette is also an actor, but is not shown on screen and is the same male voice-over in all vignettes (in a previous study the gender of the physician voiceover was included as an experimental factor but showed no consistent effects).

After viewing the vignette, physicians were asked to identify their primary diagnosis, any additional secondary diagnoses under consideration, their diagnostic certainty for all diagnoses (on a scale of 0–100), and how they would treat the patient in terms of asking for additional information, performing physical examinations, ordering tests, prescribing medications, giving lifestyle advice, and referring to other physicians. Next, an open-ended interview segment was used to explore how decisions were made, specifically in regard to the diagnostic differential and treatment trajectory. The think-aloud portion of the interview focused on two questions, with several optional probes available for interviewers to use if needed: 'Could you think back to the beginning of the consultation and tell me how you developed your differential diagnosis?' and 'Could you think back and tell me how you developed your treatment plan and what your strategy is for treating this patient?' These audio-taped segments were transcribed by a professional transcription service in batches as interviews were completed, and each transcript was one to three pages in length. These data are the focus of the present analysis.

Analytic strategy

We implemented a modified version of Strauss and Corbin's (1990) three-step approach to coding and analysis (open coding, axial coding, and selective coding) to account for a team-based approach of three coders. During open coding, transcripts were reviewed line by line by two research associates and KEL; core concepts were identified, and a preliminary coding scheme was developed and implemented on a 20 per cent subset of the data. Bi-weekly meetings were held to discuss coding inconsistencies and code criteria, and disagreements in coding were resolved by discussion until consensus was reached; extensive notes were kept on these meetings and resulting coding decisions. Two coders completed the remaining data. After coding for larger themes, the same process was used to do detailed coding to identify sub-themes and relationships between codes. The final codebook included 22 codes with detailed inclusion and exclusion criteria, examples of text segments, and other notes to assist multiple coders. The qualitative analysis software program Atlas.ti was used to facilitate the coding and management of data.

Results

The quantitative data in our study replicate results from previous studies showing CHD disparities. While the vast majority of physicians correctly identified CHD as the primary diagnosis (98.8%), diagnostic certainty (a more robust predictor of clinical actions than diagnosis alone) for CHD was varied, with an average of 57.4 on a scale of 0–100 and a standard deviation of 23.0 [unpublished manuscript]. Below, we describe the various reasoning processes leading to these different outcomes.

Pathways to CHD certainty

While a wide range of symptoms were taken into consideration, a majority of physicians focused on three core features as necessary and sufficient for their primary diagnosis to be CHD: (1) chest pain that (2) worsens with exertion and (3) is relieved with rest. Despite identical presentation of symptoms in the vignette (including non-verbal gestures), physicians varied widely in their descriptions of the location of the pain (*e.g.* ‘epigastric’, ‘substernal’, ‘midsternal’, ‘intrascapular’, ‘radiating to the back’, ‘upper abdomen’) and its character (*e.g.* ‘pain’, ‘discomfort’, ‘nonspecific chest symptoms’, ‘different from regular heartburn’, ‘vague indigestion feeling’, ‘vague bad feeling’). For some physicians, however, pain alone was sufficient for considering a CHD diagnosis:

I say anything that hurts or is not feeling right above the belly button has to be considered possibly as heart disease or coronary disease or angina (104826, male physician, more experienced).

I’m somewhat concerned always about chest pain in an individual this age certainly as being the onset of heart disease and that would seem to perhaps be the most important to make an early diagnosis on in a gentleman this age (107341, male physician, more experienced).

For most, however, the patient’s report of the pain worsening with exertion and subsiding with rest triggered the exclusion of alternative possibilities and led them to move CHD to the top of their diagnostic differentials. In their explanations of this diagnostic logic, physicians repeatedly cited an anecdote the patient told in the vignette about having the pain when walking briskly to the car in cold weather after eating dinner in a restaurant, followed by relief after resting in the car and taking some deep breaths:

So the clue [is in] the cold onset and onset by walking around. And then it went away in a few minutes by resting. So that’s angina - his fullness in the chest that goes away with rest and then it comes on with exertion or cold. Angina raised its angry head at that point and sort of stayed in my mind...(110576, male physician, more experienced).

The thing that was really urgent I thought was the walking in the cold situation and suddenly feeling really bad. That screams cardiac (134982, female physician, more experienced).

Two physicians specifically mentioned that the vignette scenario evoked an illustration they were shown in medical school to depict CHD. The picture, part of a widely-used atlas of medical illustrations (Runge and Ohman 2004) shows a grey-haired white male exiting a restaurant on a snowy night, clutching his chest (Figure 1).

The physicians commented:

I remember as a med student very well one of my professors putting up this Netter picture of this guy walking outside in the snow, out of the restaurant...He says, ‘Remember this picture, this will represent coronary artery disease to you in the future. Drill this into your head.’ And [the vignette] had all that... (112532, male physician, less experienced).

[T]his calls to mind immediately a picture in Netter's Atlas of Coronary Artery Disease (*sic*) where you see the man leaving the restaurant clutching his chest (135542, female physician, less experienced).



Figure 1 *Angina Pectoris* as depicted in *Netter's Cardiology* (2004). Netter illustration used with permission of Elsevier Inc. All rights reserved.

Together, these three classic symptoms led many physicians to a primary diagnosis of CHD. To the extent that the presentation in the vignette was seen by physicians as mirroring the examples they were taught in their medical training (e.g. male, white, older), they may have more readily closed in with certainty on a primary diagnosis and confidently relegated alternative diagnoses to secondary status. In cognitive psychology, this type of bias is referred to as a representativeness heuristic (Galanter and Patel 2005).

Particularly for female patients, however, the pathway to a CHD diagnosis is not so straightforward, as reflected in studies showing they are perpetually under-diagnosed and undertreated (Bird *et al.* 2007, Chou *et al.* 2007). This juncture in the decision-making process—wherein physicians seriously engage alternative non-CHD diagnoses—is especially critical for women as it is a point where diagnostic certainty for CHD tends to be lowered. Physicians demonstrated awareness of studies showing that women with CHD are known to present with a wider range of symptoms than their male counterparts:

...[W]ith females it's known that their presentation of a coronary artery syndrome is not typical as it would be for males with an exertional angina precipitated by exertion, cleared with rest; substernal chest pain radiating to the jaw or the left shoulder (105537, female physician, more experienced).

A lot of women are very nonspecific and vague, which is the difference of why you have to pay attention to their details as opposed to men because cardiovascular risk is the biggest cause of death in women and is usually with vague symptoms that they complain about (133562, female physician, less experienced).

However, in the vignette, women presented with precisely the same symptoms as men, all of which were specific and 'typical' rather than 'nonspecific' (as evidenced by their consistency with medical education materials in Figure 1). Still, this knowledge about possible gender differences in symptomatology seemed to contribute to physicians' decreased certainty for CHD in female patients compared to male (a cognitive pattern sometimes referred to as gender bias in psychology (Croskerry 2003)).

Diagnostic detours

Gastrointestinal conditions

Of the alternative diagnoses identified by physicians, gastrointestinal (GI) conditions were the most frequent, on a par with CHD diagnoses: 98.0% of physicians identified some type of GI diagnosis as part of their differential with average diagnostic certainty of 54.7 out of 100 (not surprising, given inclusion of the 'red herring' symptoms discussed earlier). Based on open-ended lists of diagnoses generated by the physicians, we defined GI conditions as including gastroesophageal reflux disease (GERD); gallbladder conditions; various inflammations of the GI tract such as H-pylori, pancreatitis, gastritis, or stomach ulcer; GI cancers; GI bleeding; and intestinal angina.

Most generally, the patient's food-related complaints and comparisons with earlier indigestion started physicians on a diagnostic pathway to GI diagnoses, as described by these physicians:

You have the same epigastric fullness with eating [that] would push you potentially to [a diagnosis of GERD or peptic ulcer]. The fact that you had radiation straight through the back is often a sign of esophageal spasm, which is associated with that same process. No history, a history that came on only with eating and the fact that it sometimes was relieved by antacids, as was her acid reflux that she described. The fact that she had a lot of burping and was having a lot of queasiness would again make you lean toward that diagnosis (109321, male physician, more experienced).

She's got a lot of symptoms of reflux. Spicy foods make it worse. She has heartburn. She uses antacids. In the second part of the interview she said sometimes the pain goes away with antacids. Now I didn't think that's what she said the first time. It's an ill-described pain and she seems to make a distinction between the heartburn and this other thing. She's gassy, she's got early satiety, she feels full all the time. That could certainly go along with reflux or a hiatal hernia (112633, female physician, more experienced).

Once started on the path toward a GI diagnosis, these physicians tended to interpret additional cues as consistent with their candidate diagnosis (a cognitive process referred to in psychology as confirmation bias). Physicians who continued to seriously entertain GI diagnoses further confirmed their GI suspicions by being less likely than their CHD-focused counterparts thoroughly to distinguish between the heartburn pain that was relieved by antacids and the 'new' pain symptoms that were not. In the second excerpt above, for example, the physician notes that there are two types of pain reported, but does not separate them sufficiently to trigger a CHD diagnosis. Symptoms that helped physicians rule out a GI diagnosis included the pain being (a) new, (b) different from previous indigestion, (c) not relieved by antacids, (d) triggered by stress, and (e) also triggered by large meals more than spicy food. For most doctors, however, the added information about exertional triggers, the effects of cold weather, and relief with rest shifted priorities to prioritise a cardiac diagnosis, as explained below:

...(T)hen she went on to say it's worse after a big meal, which makes you think it's probably more GI. It lasts for a few minutes, that could [imply either GI or cardiac]. When she mentioned that it was worse in cold weather that kind of raises your eyebrows and makes you think a little bit more about the possibility of heart disease (107074, female physician, more experienced).

Initially I said a more gastrointestinal...But then the more he talked and he said it was often after a large meal and then he, you know, some kind of exertion like walking quickly to the car, he's out of breath, and it was cold, or a stress, that sort of coming into the picture that that made it worse, this was starting to sound more and more cardiac (111631, female physician, less experienced).

Several critical features of the GI diagnosis facilitate its clear elimination from the differential, allowing CHD to re-emerge as the most probable diagnosis: symptoms were often seen as uniquely indicative of CHD rather than GI conditions; GI problems were seen as having clear and distinct causal relationship to the observed symptoms (thereby implying distinct treatments); and (as discussed in more detail below) physicians had various strategies to manage both potential conditions until any uncertainty was resolved through additional testing.

Mental health conditions

After GI conditions, mental health conditions were the next largest group of secondary diagnoses, with 76.6 per cent of physicians identifying at least one mental health condition as one of their diagnoses, with average diagnostic certainty of 43.1 out of 100. Based on open-ended lists of diagnoses generated by physicians, we defined mental health conditions to include depression, anxiety, panic and other mood-related disorders. The use of nonspecific and social cues, in combination with an ambiguous causal relationship between mental health and chest pain, made it difficult for patients on this pathway (disproportionately women) to be successfully routed back to a CHD diagnosis with sufficient certainty to trigger treatment.

A frequently mentioned cue was the patient's report that his/her spouse had commented on the patient 'not (being) myself lately' and being 'difficult to be around', but physicians varied in their interpretation of this information. Many viewed this comment as evidence of stressful marital conflicts, such that depression would be secondary to marital problems:

[A]t age 55 there are any number of things, even in a good marriage, that could be coming to a head....I started thinking more and more that maybe she is a woman with a lot of stress from her home situation with her husband... (120159, female physician, more experienced).

Still other physicians interpreted the comment as evidence of a lack of interest or enjoyment in daily life more generally, with depression as primary:

I get the impression that this is a new thing for him and he's not really enjoying life. He's not going anywhere or doing things, and so that is also a concern. So that's what makes me think that he may be depressed. (116174, female physician, more experienced)

Symptoms related to mental health problems were seen as nebulous in terms of their causal ordering with cardiac or GI problems. For example, pain could be brought on by stress, such that depression was driving the constellation of symptoms, including possibly cardiac problems as well:

And symptoms are pretty typical triggered by a large meal and by stress. He mentioned that when he gets stressed out, irritated by his wife's remarks, I believe, it brings up the chest pain (109092, female physician, more experienced).

Also when they said that she had pain between her shoulder blades, you see a lot of that with stress. And when she said she had decreased energy and decreased appetite, you see a lot of that with stress and depression (107846, male physician, less experienced).

For other physicians, mental health problems were secondary diagnoses and some speculated that depression in particular would be alleviated if the chest pain could be treated:

[T]he depression may actually get better if some of his physical symptoms were lightened, and so that's one reason for sort of holding off on the treatment there (117025, male physician, less experienced).

I think if we can wait until she comes back and see how she is doing and then we can talk about antidepressants. It may be that if she finds out what's wrong she may quit

being so irritable and if she feels better physically she may not be so irritable (107074, female physician, more experienced).

Other social behavioural cues reported as indicators of mental health problems included feeling irritable, having less energy and/or feeling fatigue, and feeling afraid to see the doctor. While these are relatively subtle variations in interpretation, they potentially implicate very different diagnostic assessments about the causes of depression (chemical imbalance, family problems, or chest pain) and avenues for effectively treating it (medicating the patient for the mental health problem directly, altering the home environment, or treating the CHD diagnosis and assuming mental health will improve if pain subsides). This conceptualisation is critical for two reasons: (1) women patients in the vignettes were diagnosed with mental illness more often than men (80.5% versus 72.7%); and (2) unlike GI conditions, mental health problems are difficult to eliminate from a differential diagnosis, thereby contributing to decreased certainty of CHD (and lower likelihood of CHD treatment). In psychology, the presence of such multiple alternative diagnoses may lead to increased uncertainty ('multiple alternatives bias') or 'diagnosis momentum' (Croskerry 2003), but a focus on cognitive errors alone overlooks the unfolding context of the reasoning process that leads to differential outcomes, as well as the importance of social cues.

The importance of social cues

Patient appearance

In addition to physiological symptoms, physicians also mentioned a range of socially-based cues as relevant for their decision making. As explained above, actors in the vignettes were hired and trained with the objective of standardising the appearance and behaviour of the patients portrayed across all the vignettes. This included making them as similar as possible with regard to weight, attractiveness, dress, mannerisms, affect and other physical characteristics. Still, 18 per cent of physicians mentioned patient appearance as a key influence on how they arrived at their clinical diagnoses. As one common example, physicians' characterisations of patients' weight varied widely, describing the patients as 'clearly thin', 'thin, almost frail looking', 'a little overweight', 'obese', and 'centrally obese with the big belly that you could see'. These assessments were relevant for physicians' assessments of the likelihood that a patient was suffering from CHD, as overweight is a risk factor for cardiovascular disease.

Evaluations of appearance extended beyond weight, however, and contributed directly to diagnostic reasoning for varied diagnoses:

Part of it was looking at him, looking at his face. He had as I recall the red face that you see in cardiac people, pulmonary people; that had a lot to do with it (112063, female physician, more experienced).

She certainly actually looks like someone that could have a malignancy, There's just something about her appearance that's bothersome (148436, female physician, more experienced).

Appearance could also contribute indirectly to diagnosis by providing information about potential seriousness of the problem. To the extent that cardiac patients often report having

a sense of doom prior to becoming sick, physicians were attentive to potential indicators of how serious the patient perceived the problem to be:

[The] patient was not distressed at that time...appeared healthy, did not look ill or did not look to be having significant problems. So I think from the standpoint of urgency, the degree of urgency wasn't extreme (109321, male physician, more experienced).

I heard and saw in the man's face...denial that something bad might be happening, worry that something bad might be happening. And that increased my suspicion that his symptoms are more serious and more debilitating than he's letting on (128775, male physician, less experienced).

Such use of patient appearance as a diagnostic factor amplifies the physicians' contributions to variations in CDM in at least two ways. First, despite the standardisation of presenting cases, physicians interpreted the same cues in varied ways (*e.g.* weight), and these differences were seemingly subjective and not based on any apparent medical evidence. Secondly, physicians varied widely in which characteristics they identified as salient to a range of diagnoses.

Perceived social support

Physicians were also attentive to the patients' perceived social support, with 37 per cent of physicians mentioning the patient's spouse as contributing to their clinical assessment. In the vignette, the patient mentions being out at a restaurant with his/her spouse just prior to an episode of chest pain, and that the spouse has commented that the patient is 'not taking care of [her/him]self', has 'been complaining a lot', and 'not doing anything about it'. Physician comments about the spouse revolved around the role of the spouse in: (1) generating stress or depression exacerbating the physical problem; (2) the patient's decision to come to the doctor; and (3) acting as an independent commentator on the patient's symptoms.

As with other elements of the vignette, physicians' interpretations of the spousal situation varied despite the same presentation across cases. Many physicians interpreted the spousal comments as essentially supportive and expressing 'concern' about the patient:

His wife's been concerned about his mood, basically (129614, female physician, less experienced).

She hesitated to come for three months with recurring symptoms until she was forced essentially by her husband, who's obviously concerned for her to come (105537, female physician, more experienced).

His concern was supported by his wife's concern, so I did not take it as a trivial complaint (132180, male physician, more experienced).

Yet other physicians interpreted the same cues as 'badgering' and thereby potentially aggravating the pain:

And then when his wife yanked his chain, he also got the pain back, so you're more and more certain [about a diagnosis] (116009, male physician, more experienced).

...[S]he keeps telling her husband about it, who notices the change and is noticing that she's not taking care of herself anymore, doesn't care about her appearances anymore, which goes more towards the depression...(111617, female physician, less experienced).

These differences in interpretation have key implications as they mark a difference in whether the spouse is contributing to the problem by triggering stress and subsequent pain (a conclusion that physicians described as corroborating both mental health and cardiac diagnoses) or is an objective observer trying to act in the best interests of the patient

Discussion

In a video vignette factorial experiment where the patient presentation of CHD was standardised, we used qualitative think-aloud data to identify the mental reasoning physicians used in their diagnostic decision making. Quantitative results from the factorial experiment show that the vast majority of physicians identified CHD as a primary diagnosis, but were relatively uncertain of that diagnosis and simultaneously entertained several alternatives. We identify which cues tend to lead physicians to CHD diagnoses versus diagnostic detours involving gastrointestinal and mental health conditions; a range of social and physiological cues that are deemed relevant for diagnostic and treatment decisions; and how these and cognitive biases build on one another. Specifically, we identify three key junctures where patients' likelihood of receiving a high certainty CHD diagnosis (and therefore CHD treatment) are facilitated or jeopardised: (1) at the point where physicians seriously entertain alternative, non-CHD diagnoses, which lowers certainty for CHD and disproportionately affects women; (2) the ruling out of a GI diagnosis, which is seen as having distinct etiology from CHD and therefore can be eliminated from a differential diagnosis with certainty or co-managed until testing provides sufficient information; and (3) the difficulty in ruling out mental health conditions, which are evidenced by nonspecific social and behavioural cues and have ambiguous temporal and causal ordering with CHD, all of which hinder their elimination from a differential diagnosis and complicate the possible management of both conditions simultaneously.

Despite a proliferation of research on health disparities and biases in clinical decision making, there is a dearth of data that allow for the identification of cognitive pathways physicians take to the diagnosis and treatment of CHD. In terms of conceptualising physician decision making, our analysis shows how, rather than just ranking candidate diagnoses into a static list according to certainty levels or potential seriousness, physicians are also concerned with how the conditions interact with one another and how they are causally and temporally ordered. While psychological work has made important contributions to our understanding of cognitive biases, a broader sociological analysis shows how social, psychological, and biological systems intersect in critical ways. Furthermore, most qualitative work addressing these types of concerns does not control for the presentation of symptoms, so it is often impossible to determine whether observed variation results from differences in patients' presentation or in physicians' responses to them. As a result, this study fills important theoretical and methodological gaps in our knowledge about some of the reasons for observed aggregate differences in clinical decision making.

To address external validity and inevitable challenges inherent in trying to replicate everyday behaviour within an experimental setting, we took four precautionary steps: first, ensuring clinical authenticity of the videotape presentation by basing scripts on clinical

experience and using professional actors/actresses (similar studies have used this approach successfully) (Feldman *et al.* 1997); secondly, asking physicians to compare the presentation of the patient in the vignette to patients they encounter in everyday practice (89.8% considered them either very typical or reasonably typical); thirdly, having physicians view the tapes in the context of their daily practice in between patient appointments; and finally, specifically instructing physicians to view the patient as one of their own and to respond as they would typically in their practice. In a direct comparison of vignettes, standardised patients, and chart abstraction, Peabody and colleagues (2000) found that vignettes were a valid and comprehensive method for measuring quality of outpatient care. Additional studies comparing vignettes with standardised patients and other methods corroborate the result that vignettes are ecologically valid for studies of medical decision making (Dresselhaus *et al.* 2000, Veloski *et al.* 2005), and medical educators have successfully used vignettes to mimic cognitive diagnostic processes in live clinical interactions (Mazor *et al.* 2007, Ogden, Edwards, and Stricken 2003).

Future research could expand several aspects of this work to build on its limitations and expand its scope. For example, despite efforts to protect external validity, a vignette-based exercise is not the same as in-person clinical treatment, and so future work could examine how our results compare with live patient presentation. Similarly, we would expect results to differ across medical conditions and different types of providers. At the same time, even with different specific study details, future research may corroborate our general results showing that diagnostic differences are less the result of any single type of cognitive or social bias, but rather the cumulative, compounding, multi-level effect of many smaller mechanisms as part of a larger 'durable narrative' linking certain types of patients with certain types of outcomes (Lutfey and Freese 2005).

This analysis underscores the role of physicians in potentially amplifying health disparities and contributing to epidemiologic rates of disease. To the extent that physician-generated diagnostic labels contribute to the compilation of official base rates, the types of inaccuracies or biases outlined above are reified — not only as they are published and made 'official', but also as physicians use Bayesian decision-making models and continue to draw on base rates in future decision making. Particularly entrenched conceptualisations, such as the long-held notion that women suffer from CHD at much lower rates than men (McKinlay, Potter and Feldman 1996), are reified as part of medical education (as in Figure 1).

The implications of this work are also relevant for researchers concerned with the inverse relationship between socioeconomic status (SES) and health. Despite the diversity of the topics and theoretical perspectives represented in this body of research on health disparities, many of these studies share a reliance on official statistics for understanding rates of disease, how they change over time, and how health is related to SES and other factors. While not all these rates are derived from diagnostic labels applied during patient-physician interactions of the type addressed in this study, many are related to these interactions. For the American Heart Association, official statistics about CHD are compiled based on well-known health surveys such as the Behavioral Risk Factor Surveillance Survey (BRFSS), the Medical Expenditure Panel Survey (MEPS), National Health and Nutrition Examination Survey (NHANES), National Health Interview Survey (NHIS), National Hospital Discharge Survey (NHDS), and others (Lloyd-Jones *et al.* 2009). For several of these surveys, individual cases become part of epidemiologic rates if the reimbursement charges indicate CHD diagnosis or treatment, or patients report ever having been told by a healthcare provider that they have CHD — either of which is related to the sorts of patient-physician encounters studied here.

Following on from Freidson's observation that 'the idea of illness may or may not have a foundation in biological reality...it always has a foundation in social reality' (1970: 212), we take as problematic social consensus surrounding the 'social fact' of CHD disparities. As suggested in a recent published exchange on the topic of naming and framing of disease (Aronowitz 2008), such framing mechanisms are sometimes acknowledged in current social science health research, but 'relegated to a massive list of 'confounders and biases' that researchers busy themselves tidying or...find ways to 'control'' (Fullwiley 2008: 15). Considered together, this work shows how the labelling of a case of CHD has implications at several levels, including not only immediate diagnostic perceptions about certain types of patients, but extending also decision making more generally, as well as epistemology of epidemiological rates and their role in ongoing health research.

Address for correspondence: Karen E. Lutfey, New England Research Institutes, 9 Galen Street, Watertown, Massachusetts 02145, United States
e-mail: KLutfey@neriscience.com

Acknowledgements

We appreciate clinical input on vignette development from Timothy Guiney MD, Alan Goroll MD, Theodore Stern MD, and John Stoeckle MD (Massachusetts General Hospital, Harvard Medical School, Boston). Generous coding assistance was provided by Megan Renfrew and Amy O'Donnell. This research is supported by NIH grant HL079174 (John B. McKinlay, PI).

Note

1 In the US, internal medicine is similar to family medicine in that it includes primary care and basic healthcare management of all areas of a patient's health. Internal medicine differs slightly from family medicine in that it typically does not include pediatrics or obstetrics, but may include more sub-specialties such as endocrinology or gastroenterology. However, physicians who reported practising primary patient care for less than 50 per cent of their time were excluded from the study.

References

- Adams, A., Buckingham, C.D., Arber, S., McKinlay, J.B., Marceau, L.D. and Link, C.L. (2006) The influence of patient age on clinical decision-making about Coronary Heart Disease in the US and the UK, *Ageing and Society*, 26, 303–21.
- Anglin, D.M., Link, B.G. and Phelan, J.C. (2006) Racial differences in stigmatizing attitudes toward people with mental illness, *Psychiatric Services*, 57, 857–62.
- Arber, S., McKinlay, J.B., Adams, A., Marceau, L.D., Link, C.L. and O'Donnell, A.B. (2004) Influence of patient characteristics on doctors' questioning and lifestyle advice for Coronary Heart Disease: a UK/US video experiment, *British Journal of General Practice*, 54, 673–8.
- Armstrong, D.L., Strogatz, D. and Wang, R. (2004) United States coronary mortality trends and community services associated with occupational structure, among blacks and whites, 1984–1998, *Social Science and Medicine*, 58, 2349–61.
- Aronowitz, R. (2008) Framing disease: an underappreciated mechanism for the social patterning of health, *Social Science and Medicine*, 67, 1–9.

- Barnhart, J.M. and Wassertheil-Smoller, S. (2006) The effect of race/ethnicity, sex, and social circumstances on coronary revascularization preferences: a vignette comparison, *Cardiology Reviews*, 14, 215–22.
- Bird, C.E., Fremont, A.M., Bierman, A.S., Wickstrom, S., Shah, M., Rector, T., Horstman, T. and Escarce, J.J. (2007) Does quality of care for cardiovascular disease and diabetes differ by gender for enrollees in managed care plans?, *Womens Health Issues*, 17, 131–8.
- Bittner, V. (2008) Angina pectoris: reversal of the gender gap, *Circulation*, 117, 1505–7.
- Bosk, C. (1979) *Forgive and Remember: Managing Medical Failure*. Chicago: University of Chicago Press.
- Britt, H., Bhasale, A., Miles, D.A., Meza, A., Sayer, G.P. and Angelis, M. (1996) The sex of the general practitioner: a comparison of characteristics, patients, and medical conditions managed, *Medical Care*, 34, 403–15.
- Chou, A.F., Brown, A.F., Jensen, R.E., Shih, S., Pawlson, G. and Scholle, S.H. (2007) Gender and racial disparities in the management of diabetes mellitus among Medicare patients, *Womens Health Issues*, 17, 150–61.
- Croskerry, P. (2003) The importance of cognitive errors in diagnosis and strategies to minimize them, *Academic Medicine*, 78, 775–80.
- Curoe, A., Kralewski, J. and Kaissi, A. (2003) Assessing the cultures of medical group practices, *Journal of the American Board of Family Practice*, 16, 394–8.
- Dresselhaus, T.R., Peabody, J.W., Lee, M., Wang, M.M. and Luck, J. (2000) Measuring compliance with preventive care guidelines: standardized patients, clinical vignettes, and the medical record, *Journal of General Internal Medicine*, 15, 782–8.
- Feldman, H.A., McKinlay, J.B., Potter, D.A., Freund, K.M., Burns, R.B., Moskowitz, M.A. and Kasten, L.E. (1997) Nonmedical influences on medical decision making: an experimental technique using videotapes, factorial design, and survey sampling, *Health Services Research*, 32, 343–66.
- Forrest, C.B., Nutting, P.A., von Schrader, S., Rohde, C. and Starfield, B. (2006) Primary care physician specialty referral decision making: patient, physician, and health care system determinants, *Medical Decision Making*, 26, 76–85.
- Fox, R. (1957) Training for Uncertainty. In Merton, R.K., Reader, G. and Kendall, P. (eds) *The Student-Physician*. Cambridge: Harvard University Press.
- Freund, K.M., Moskowitz, M.A., Lin, T.H. and McKinlay, J.B. (2003) Early antidepressant therapy for elderly patients, *American Journal of Medicine*, 114, 15–9.
- Galanter, C.A. and Patel, V.M. (2005) Medical decision making: a selective review for child psychiatrists and psychologists, *Journal of Child Psychology and Psychiatry*, 46, 675–89.
- Goodman, D.C. and Grumbach, K. (2008) Does having more physicians lead to better health system performance? *Journal of the American Medical Association*, 299, 335–7.
- Gurjeva, O.S., Bukhman, G., Murphy, S. and Cannon, C.P. (2005) Treatment and outcomes of eastern Europeans with coronary syndromes in OPUS-TIMI 16, *International Journal of Cardiology*, 100, 101–7.
- Krupat, E., Irish, J.T., Kasten, L.E., Freund, K.M., Burns, R.B., Moskowitz, M.A. and McKinlay, J.B. (1999) Patient assertiveness and physician decision-making among older breast cancer patients, *Social Science and Medicine*, 49, 449–57.
- Lloyd-Jones, D., Adams, R., Carnethon, M., et al. (2009) Heart disease and stroke statistics - 2009 update: a report from the American Heart Association Statistics Committee and Stroke Statistics Subcommittee, *Circulation*, 119, e21–181.
- Lutfey, K.E. and Freese, J. (2005) Toward some fundamentals of fundamental causality: socioeconomic status and health in the routine clinic visit for diabetes, *American Journal of Sociology*, 110, 1326–72.
- Lutfey, K.E., Link, C.L., Grant, R.W., Marceau, L.D. and McKinlay, J.B. (2009) 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.

- Maly, R.C., Umezawa, Y., Ratliff, C.T. and Leake, B. (2006) Racial/ethnic group differences in treatment decision-making and treatment received among older breast carcinoma patients, *Cancer*, 106, 957–65.
- Mazor, K.M., Haley, H.L., Sullivan, K. and Quirk, M.E. (2007) The video-based test of communication skills: description, development, and preliminary findings, *Teaching and Learning in Medicine*, 19, 162–7.
- McKinlay, J.B., Potter, D.A. and Feldman, H.A. (1996) Non-medical influences on medical decision-making, *Social Science and Medicine*, 42, 769–76.
- McKinlay, J.B., Lin, T., Freund, K.M. and Moskowitz, M. (2002) The unexpected influence of physician attributes on clinical decisions: results of an experiment, *Journal of Health and Social Behavior*, 43, 92–106.
- McKinlay, J.B., Link, C.L., Freund, K.M., Marceau, L.D., O'Donnell, A.B. and Lutfey, K.E. (2007) Sources of variation in physician adherence with clinical guidelines: results from a factorial experiment, *Journal of General Internal Medicine*, 22, 289–96.
- Ogden, P.E., Edwards, J.C. and Stricken, A.G. (2003) Portable interactive faculty development for multiple site clinics using computer-based CD-ROM modules, *Medical Education*, 37, 1043–4.
- Paget, M.A. (1988) *The Unity of Mistakes: A Phenomenological Interpretation of Medical Work*. Philadelphia: Temple University Press.
- Peabody, J.W., Luck, J., Glassman, P., Dresselhaus, T.R. and Lee, M. (2000) Comparison of vignettes, standardized patients, and chart abstraction: a prospective validation study of three methods for measuring quality, *Journal of the American Medical Association*, 283, 1715–22.
- Popescu, I., Vaughan-Sarrazin, M.S. and Rosenthal, G.E. (2007) Differences in mortality and use of revascularization in black and white patients with acute MI admitted to hospitals with and without revascularization services, *Journal of the American Medical Association*, 297, 2489–95.
- Runge, M.S. and Ohman, M.E. (2004). *Netter's Cardiology*. Philadelphia: Saunders.
- Schulman, K.A., Berlin, J.A., Harless, W., et al. (1999) The effect of race and sex on physicians' recommendations for cardiac catheterization, *New England Journal of Medicine*, 340, 618–26.
- Starfield, B., Forrest, C.B., Nutting, P.A. and von Schrader, S. (2002) Variability in physician referral decisions, *Journal of the American Board of Family Practice*, 15, 473–80.
- Strauss, A. and Corbin, J. (1990) *Basics of Qualitative Research: Grounded Theory Procedures and Techniques*. Newbury Park, CA: Sage.
- Veloski, J., Tai, S., Evans, A.S. and Nash, D.B. (2005) Clinical vignette-based surveys: a tool for assessing physician practice variation, *American Journal of Medical Quality*, 20, 151–7.
- von dem Knesebeck, O., Bönke, M., Siegrist, J., Marceau, L., Link, C., Arber, S., Adams, A. and McKinlay, J.B. (2008) Country differences in the diagnosis and management of coronary heart disease - a comparison between the US, the UK and Germany, *BMC Health Services Research*, 8, 198.
- Weisz, D., Gusmano, M.K. and Rodwin, V.G. (2004) Gender and the treatment of heart disease in older persons in the United States, France, and England: a comparative, population-based view of a clinical phenomenon, *Gender Medicine*, 1, 29–40.