Introduction

This is a book about the processes by which we recognize, name, classify, and find meaning in illness. It focuses on the ways that twentieth-century U.S. investigators, clinicians, and patients have come to recognize a new disease or disease etiology, give it a name, place it in a certain class of diseases or causes, and give it individual and social meaning. My working assumption is that a new consensus about illness is usually reached as a result of negotiations among the different parties with a stake in the outcome. Insights from the clinic and laboratory create options for a new disease category or a different meaning of an existing name, but do not ultimately determine the outcome of a largely *social* process of negotiation.

Detailed knowledge of these negotiations matter because their outcomes matter. The recognition, naming, and classification of disease is central to so many aspects of late-twentieth-century life, whether we are a patient receiving a diagnosis to explain painful and frightening symptoms, a researcher conducting a clinical trial, a worker claiming disability, or an advocacy group pressing the government to investigate an apparent outbreak of a previously undescribed illness. Yet the processes by which we decide what is a disease, what types of suffering remain nameless and invalid, and what names, causes, and meanings we attach to different types of suffering are generally taken for granted.

My approach to understanding these processes is historical and contextual. The core of this book are six case studies of twentieth-century illness-in-flux, situations in which a new disease, etiology, and/or classification type was recognized and named. We may now know the ultimate outcome of these changes, but the choices that were available to investigators, clinicians, and patients and the processes by which change occurred are not generally appreciated.

By making these choices and processes explicit, I hope to contribute to a better understanding of the fundamental terms of many contemporary

clinical and policy controversies. The contemporary cultural and medical landscape is littered with controversial disease entities, public policy debates that hinge on definitions of disease and disability, and angst about the reigning biomedical model of disease. While all of these stress points hinge on definitions, classifications, and meanings of ill health, the underlying historical context for contemporary controversies is frequently ignored. In my view, it makes little sense to argue whether such and such disease is legitimate without an understanding of what we generally mean by "legitimate" disease; to argue whether government entitlements should cover disease X, without understanding how and why particular categories of ill health are grouped and named together and granted special status; or even to criticize our health care system as dehumanized and reductionist, without some understanding of the historically conditioned values and interests that have framed the basic building blocks of that health care system.

I will now introduce some general themes and suggest the clinical relevance of my approach by briefly relating a few experiences of friends, family, and patients. Each of these vignettes emphasizes a different way that underlying and often unresolved issues about definitions, classifications, and meanings of disease influence contemporary medical encounters.[1]

Case Studies

Harold: The Individual as the Cause of Disease

One of my first patients as a third-year medical student was Harold, an 18-year-old with Crohn's disease who was admitted to the surgical service for repair of abnormal connections between his inflamed bowel and surrounding organs. After his operation, he developed an acute psychosis that required a stay in the neuropsychiatric evaluation unit. No satisfying answer for his psychosis was ever made, but I got the impression from the house staff and the gastroenterologists that it was not surprising that a Crohn's disease patient would have psychiatric problems. Such beliefs are common, discussed among doctors but rarely mentioned in contemporary textbooks and review articles.

Following a suggestion made by Harold's consulting gastroenterologist, I reviewed the older medical literature on ulcerative colitis, which many believe to represent a similar, if not identical, pathophysiological process as Crohn's disease and which is generally grouped with Crohn's

disease under the label "inflammatory bowel disease." I learned that in the 1930s and 1940s, ulcerative colitis was generally felt to be a psychosomatic disease by most physicians and was treated as such in medical texts.

Although what was meant by psychosomatic disease was not entirely clear, I learned that physicians in that era tried to answer questions about Harold's type of chronic disease that were not even being asked, let alone answered, by my teachers and contemporary medical texts: Why was he rather than someone else afflicted with this disease? Why was he having an exacerbation now? Did his Crohn's disease cause his psychosis or was his psychosis a cause of his disease? The fact that speculations about the answers to these questions are today whispered at the bedside but not discussed in the medical literature or formal case presentations is curious and suggests that our formal systems of medical discourse systematically exclude certain categories of knowledge and speculation. Why do these gaps exist? How has mainstream medicine in other eras accommodated persistent concerns about the relationships among social and psychological factors, individual predisposition, and the cause, appearance, and course of disease?

Elizabeth: Disease or Personal Diagnosis?

Elizabeth consulted a general internist, her first such visit in years, because she had gradually developed abdominal pain over the preceding few weeks. After a complete history and physical, her internist felt that the two leading diagnostic possibilities were peptic ulcer disease and pain from gallstones. He planned to get an ultrasound of Elizabeth's abdomen to look for gallstones at some future date if the pain was not relieved with antacids. This was the internist's standard approach to a patient with mild abdominal pain who he did not suspect had a serious, acute disease. Often such patients would get better or at least never return to the office. Those who returned without much relief would get more diagnostic tests and perhaps more specific medications.

A few weeks after the initial consultation, however, Elizabeth's pain got much worse and was accompanied by fever. She was away from home so she went to the closest emergency room, where the on-call surgeon made a diagnosis of an obstructed gallbladder, took her to the operating room, and removed it.

It is possible that Elizabeth's gallbladder disease might have been diagnosed earlier if her internist had placed more emphasis on Elizabeth rather than used a standard approach to the average patient. Had he

tried to assess her pattern of symptom recognition and threshold for seeking medical care, he might have suspected that Elizabeth had a problem that required more urgent attention. He would have learned, for example, that Elizabeth had almost given birth to her first child on the way to the hospital because she did not believe that her labor pains were severe enough to signify the late stages of labor.

Why is it that physicians do not routinely elicit this type of information from patients? Diagnostic models based on disease as a purely biological entity and the "average" individual generally exclude knowledge about individual differences and social factors. Knowledge and approaches that might allow physicians to learn what is best for individual patients, often based more on observation and listening than on "objective" data, are both undeveloped and undervalued. How has this situation developed? What have been the obstacles to expanding our clinical gaze?

Margaret: Disease or Personal Prognosis?

Margaret was 83 years old and in good health when she became jaundiced, itchy, and tired over the course of a few weeks. She was admitted to a local hospital for an ERCP (endoscopic retrograde cholangiopancreatogram), a procedure that her gastroenterologist thought might result in a diagnosis to explain her jaundice and possibly relieve it. During the procedure, Margaret's gastroenterologist biopsied a mass in her pancreas that turned out to be a cancerous tumor. He told Margaret and her family that she had only a few months to live. More than two years later and after numerous procedures to bypass her clogged biliary system, she was still alive, defying the textbook odds. In retrospect, it is possible that the diagnosis of pancreatic cancer, while correct, was something of an incidental finding. Her presenting symptom of "painless jaundice" was probably not due to cancer but to another process that narrowed her bile duct. A year prior to the cancer diagnosis, Margaret had a benign stricture in her common bile duct that required dilatation.

The textbook prognosis for pancreatic cancer, itself only an educated guess under the best circumstances, might have been inaccurate in Margaret's case because it was primarily derived from the clinical experience of patients whose symptoms led to the cancer diagnosis. In general, prognostic schemes do not incorporate the circumstances in which disease is first diagnosed, despite their logical and clinically evident importance. Why have they not? Should they? Why have research and clinical practices focused so narrowly on localized pathology rather than the individual who suffers disease?

Marty: Physician's or Patient's Definition of Disease?

Marty, a 50-year-old business executive who smokes and has mildly elevated blood pressure, had been taking nitroglycerine and other medications for angina pectoris for the past three years. He described his chest pain alternatively as indigestion and a pressure, often related to exertion but also occurring after meals. The pain was sometimes relieved with rest but more often by burping. After a particularly severe episode, he was admitted to the hospital to "rule out myocardial infarction," that is, heart attack. He did not have a heart attack and subsequently underwent coronary angiography, which revealed "clean" (normal) coronary arteries. He was told that he did not have angina pectoris and was discharged.

In an earlier era, no test could have taken away the diagnosis of angina pectoris from such a patient – the symptoms defined the disease. What is Marty suffering from? Would he be better served by the older definition? Changing disease definitions often reflect compelling beliefs in transition and/or conflict. In this case, a priority given to the way a disease is experienced by a patient is in conflict with the belief that a specific, measurable, and visible anatomic abnormality is the best way to define disease. What are the stakes in these underlying conflicts? How have physicians, patients, and others sought to resolve or moderate them? What have been the consequences of particular solutions?

Larry: Who Has the Authority to Define the Scope of Disease?

Larry, a 50-year-old man, came to my office with a Hickman catheter (a semipermanent intravenous line that surgeons usually place in cancer patients who are expected to undergo long-term chemotherapy) for a second opinion about the management of his Lyme disease. He had been to many doctors searching for an explanation of, and treatment for, long-standing fatigue and muscle aches. A physician at a self-styled Lyme disease "center" diagnosed him as having chronic Lyme disease. After initial antibiotic therapy failed to improve his symptoms, the Hickman catheter was placed to deliver repeated courses of intravenous antibiotics. Larry began to doubt the wisdom of this treatment course, although he had been initially relieved when physicians had diagnosed a real and treatable disease. Neither Larry's Lyme disease diagnosis nor his treatment conformed to the recommendations of Lyme disease experts. When cases similar to this were investigated by the Centers for Disease Control (CDC) after a cluster of antibiotic-related gallbladder disease was re-

ported to state health authorities, the offending physicians defended their practices by questioning the very authority of Lyme disease experts to define what was the best criterion for diagnosis and treatment.[2] Many individual patients and Lyme disease groups support their position, arguing, for example, that the narrow clinical criterion suggested by experts excludes many patients who are really suffering from Lyme disease.

Disease definition has increasingly become a publicly debated issue. Formerly such controversy was reserved for borderland medical diagnoses – alcoholism or homosexuality. Now debates rage even in "legitimate" diagnoses such as Lyme disease and AIDS. How did the diagnosis and scope of Lyme disease and other chronic diseases become so controversial? What are the stakes? Who are the likely winners and losers?

Louis: Living or Dying with a Diagnosis?

Louis is a successful engineer, still working in his 70s, who thought it might be a good idea to request a blood test for prostate cancer after the same test led to the diagnosis and surgical treatment of cancer in one of his friends. I explained that using this test to screen for prostate cancer was controversial, but Louis thought it was nevertheless a good idea to get it done. When I received notice that the blood test was positive for prostate cancer, I referred Louis to a urologist, who did a prostatic ultrasound looking for cancer. No cancer could be seen. The urologist then did six "blind" biopsies of Louis's prostate gland, the last of which contained a small focus of cancer.

In Louis's engineering work, he was something of a decision analyst, so he asked piercing questions to the urologist about the three options presented to him: watchful waiting (doing nothing), radiation therapy, or surgical removal of his prostate gland. Louis learned that although no one knew what the odds were that he would ever experience symptoms or die from a small cancer that was picked up by blind biopsy after a positive prostate cancer blood test, the chances that he would die from prostate cancer were small. He kept returning to his own situation – what if the surgeon had done only five rather than six biopsies? I tried to help Louis with his decision by telling him about autopsy studies that have found small prostate cancers in about half the men his age who died of unrelated reasons and that there was no – as yet – good evidence that people who underwent surgery were better off. And his cancer might have even less potential for harm than most because of its small size and almost fortuitous discovery.

Although the surgeon felt that the best course was either radiation

therapy or surgery, Louis could not convince himself that the high probability of side effects such as impotence and urinary incontinence were less problematic than living with the small possibility of dying from prostate cancer. Upon hearing his initial decision to do nothing about his prostate cancer, Louis's family's reaction – born out of a very understandable concern over a seemingly irrational, self-destructive decision – was to accuse him (with me as accomplice) of trying to kill himself and ruin their family life. Louis eventually changed his mind, opted to receive radiation therapy, and continues in good health.

It was more than the dearth of accurate statistics about prostate cancer – its natural history and treatments – that made this decision so difficult for Louis and the many people who care about and for him. Perhaps most problematic were the meaning and connotations of the word "cancer." Even if one could soberly weigh the "oranges" of the cancer's probable harm against the "apples" of the treatments' side effects, the thought of living with a cancer – especially one that could be gotten rid of – is very troubling for most people. It goes against the underlying meaning of cancer to believe its probability for causing serious morbidity and mortality are very low. We expect an inexorable progression of cancer from bad to worse. What does cancer mean? When is a cancer's potential for harm so low that it might no longer be considered a cancer? Should cancer be defined and diagnosed by pathologists or others? What factors have led to the proliferation of diseases – such as Louis's "stage" of prostate cancer – whose primary meaning lies in their statistical risk rather than the symptoms they cause?

The Problem of Idiosyncrasy

Each of these friends and patients faced a problem that in part resulted from the conventional ways we attribute cause (Harold), make diagnoses (Elizabeth), determine prognosis (Margaret), label suffering (Marty), and define (Larry) and find meaning in (Louis) disease. In particular, these patients' experiences suggest that a major area of disagreement at the root of many clinical and policy controversies concerns the ways we accommodate what individuals bring to disease, what I shall generally refer to as the problem of "individual idiosyncrasy," or just idiosyncrasy. Should we have prognostic models that incorporate information about the way individuals present to the medical system? Should the doctor–patient encounter always involve a personal, not just a disease, diagnosis? Should the patient's experience or the doctor's test define disease? Can we have models of disease etiology that incorporate the multitude

of social and psychological factors that shape the appearance of disease in the individual?

In the chapters that follow, I explore how medical and lay persons responded to questions about idiosyncrasy that mattered to them but have been typically outside the biomedical purview. The key questions often involve individual predisposition and responsibility for disease. The answers to the "why me?" and "why now?" questions about disease, especially chronic disease, have varied according to the character of the specific biologic processes involved, the disease spectrum of an era, and the persons who are answering the questions. By sampling diseases that are from different time periods and that represent different biological characteristics, I demonstrate the pervasive, if largely hidden, influence of these underlying questions on biomedical and epidemiological investigations, disease definition, clinical practices, health policy, and – ultimately – the patient's experience of illness.

In trying to understand clearly the available strategies for dealing with the question of individual idiosyncrasy, I have found it helpful to view patients and doctors as having to continually negotiate between two competing ideal-typical notions of ill health – illness as "specific disease" and illness as "individual sickness." Historians of medicine have labeled as "ontological" the view that diseases are specific entities that unfold in characteristic ways in the typical person. In this framework, diseases exist in some platonic sense outside their manifestations in a particular individual. The other compelling account of illness, the "physiological" or "holistic," stresses the individual and his or her adaptation, both psychological and physical, to a changing environment. In this framework, illness exists only in individuals. These ideal-typical notions have been in a state of dynamic tension since antiquity.[3]

It might be argued that with the ascendancy of the germ theory of disease in the late nineteenth century, the ontologic view of illness gained a lasting preeminence. When it was discovered that particular microorganisms caused distinctive pathological derangements and clinical presentations, this etiology became the prototype for explaining most diseases and sickness in general, up to and including the credo of contemporary molecular biology: one gene, one protein, one disease. Individual factors such as the role of emotions, lifestyle, and social class in the etiology, appearance, course, and distribution of disease were, in the course of the twentieth century, relegated to the margins of medical and lay concerns.

While not without merit, this monolithic view of changing "disease theory" simplifies the continual negotiation and shifting balance in medical research, clinical practice, and social thought between ontological

and holistic orientations. In the twentieth century, the appeal of ontological models of disease, even in the domain of acute, infectious disease, has been tempered by questions about individual predisposition to disease and the social, nonspecific basis of dramatic trends in disease morbidity and mortality.

For example, epidemiologists and others recognized early in this century that in a polio epidemic a significant percentage of the population may become infected, yet only a small fraction develops symptoms and an even smaller fraction, paralysis. What principles might explain individual predisposition to clinically apparent disease? Similarly, Renee Dubos, Thomas McKeown, and others have emphasized that the historical decline of tuberculosis mortality in this century has been a constant one, seemingly uninfluenced by the introduction of specific public health approaches and clinical interventions such as isolating infectious individuals and treating with antibiotics the silently infected, as well as individuals with clinically apparent disease.[4] Such measures seem to be mere epiphenomena, the important determinants of declining tuberculosis mortality residing elsewhere, for example, in improved nutrition and economic and technological development generally.

Moreover, in any particular time and setting the balance of ontological and holistic views of illness will be dependent on the existing disease burden. Oswei Temkin noted, for example, that Thomas Sydenham, the arch-ontologist, "lived at the time of the great plague of London, and the plague, I understand, has little concern with individual variations."[5] We live in an era in which much of our health care expenditures and illness experience are due to chronic disease. We accommodate the holistic view when we acknowledge that much of the suffering in chronic disease is not amenable to "magic bullets" and is highly dependent on individual and social circumstances. Even at the level of biological understanding, chronic disease raises the visibility of the individual dimension because etiologic models generally assume that multiple environmental and genetic factors operating at the level of the individual organism have been interacting for long periods of time before the onset of overt disease. AIDS is exemplary, since the identification of the organism and modes of transmission have not by themselves led directly to a basic understanding of either the disease's pathogenesis or an effective treatment – not to mention the ways that disease originates and develops in individuals.

Although I will frequently refer to the dichotomy between ontological and holistic ideal-typical notions of illness to make sense of underlying tensions in disease recognition, naming, and categorization, I do not wish to make this a dominant idea that structures all aspects of my

historical analysis. For one thing, as we shall see in later chapters, despite the rhetoric of a self-conscious and oppositional holism, or an insistent biomedical reductionism, the actual research and practice of investigators and clinicians inevitably take on many features of their opposites. Witness the ironies of contemporary efforts to find the active molecular building blocks of holistic therapies in order to manipulate them in randomized controlled trials under the auspices of a newly created National Institute of Health program devoted to studying holistic medicine. In the career trajectory of a holistic critic of medicine seeking academic success or the clinical experience of a late-twentieth-century medical specialist taking care of an individual whose pain does not neatly fit any available disease category, there is an ineluctable fusing of perspectives. This should not be surprising since these underlying notions are, as Charles Rosenberg put it in another context, "mutually constitutive" in medical practice.[6]

Nor am I entirely happy with the connotations of terms such as "holism" and "ontology," "individual sickness" and "specific disease." In some situations, the contrast between universality and idiosyncrasy might be more evocative of the underlying tension without carrying the historical baggage that a term such as "holism" evokes. But the important danger to avoid is that of reifying this or any of the other related dichotomies. There is no self-evident boundary between the specific, objective, and pathological, on the one hand, and the holistic, subjective, and experiential, on the other. The distinction is necessarily an oversimplification of a more complex and nuanced reality in which elements of both ways of thinking about and perceiving disease are present. While this dualistic view may serve as a useful way to assign professional roles and spheres of investigation, or even to approach moral issues (e.g., attributing responsibility for disease), it can potentially weaken the position of those advocating a more patient-centered system of medical care by helping to uphold an artificial boundary between the science and the "art" of medicine.[7]

Making Sense of Illness: Interactions among Social and Biologic Determinants of Disease Meaning

In order to characterize how twentieth-century investigators, clinicians, and patients have recognized and agreed upon a new disease label, category, or cause, I frequently employ the term "the social construction of disease." This term is generally used to describe historical and other approaches that analyze and describe the interaction between social