CHAPTER TWO

Histories of the Human Subject

An oft-repeated claim from the mid-1980s onward is that the field of medicine has long presumed a "male norm" and that various groups, especially women, have been invisible to researchers and clinicians.[1] Critics of the underrepresentation of women, children, the elderly, and racial and ethnic minorities as research subjects have suggested that privileged white males in their twenties through their fifties were, for too long, taken to be the "standard human"—the reference point from which knowledge about human health and illness flowed. This was precisely Bernadine Healy's argument, as described in the introduction: the former NIH director blasted "the orthodoxy of sameness and the orthodoxy of the mean" and described how the 35-year-old white male had long served as the "normative standard" in medical research.[2] Or to quote the blunt assessment of an "outreach notebook" distributed by the NIH in 2002 to assist researchers with the recruitment and retention of women and minority research subjects: "Historically, the typical and usual research participant was a white male."[3]

Like any influential representation of history, this one merits scrutiny. Have other groups besides middle-aged white men simply been overlooked by medicine? To the degree that they have been overlooked, was it because they were seen as so similar to the "standard human" that specific study wasn't deemed necessary? Or was it because they were seen as fundamentally different yet socially inferior—in which case specific study was no one's priority?[4] And to the degree that groups such as women and racial and ethnic minorities *have not* been overlooked and instead have been the object of explicit study, what have been the motivations for such studies, and what generalizations about humanity have been drawn from them?

In this chapter, I lay the groundwork for understanding recent changes in clinical research practices by investigating how medical researchers at earlier historical moments decided on which people to study. Addressing this issue demands attention to three crucial sets of oppositions that surface in the history of medical experimentation:

- *Privilege versus vulnerability*: Under what historical conditions do socially advantaged groups become the objects of researchers' attention? Conversely, when do researchers find it more convenient to subject marginalized, vulnerable, or captive populations to the risks inherent in medical experimenting?
- *Sameness versus difference*: From a medical standpoint, how alike or unlike are different social groups imagined to be?
- *Generalizability versus specificity*: When do researchers believe that results from experiments conducted on one group can be extrapolated to other groups, and when are they hesitant to make such generalizations?[5]

The basic claim of this chapter is that there is no single set of answers to these questions that universally captures or characterizes Western medical thought or practice in recent centuries or even decades. Using examples and arguments drawn primarily from the work of historians, I will show how medical experts have arrived at varied responses to these questions at different historical moments, and I will demonstrate that, even at a given point in time, researchers often have not been unanimous or consistent in their views. Therefore, it is important to examine the distinctive and sometimes idiosyncratic ways in which medical authorities have resolved these different debates. But in order to do so, we need to sketch the outlines of several intertwining histories—of medical philosophies, methodologies of testing, and understandings of the ethics of experimentation on humans. As I review these histories, I will consider the implications of each for the representation, underrepresentation, or exclusion of various groups from medical research.

What will this review of the historical record allow us to conclude about reformers' claims concerning inclusion and the standard human? Were the advocates of social change accurate in their portrayal of history, or have they gotten the history wrong? The question is difficult to answer for several reasons, not least because the rhetoric of the movement in favor of inclusion is rather imprecise about *when* these claims are meant to apply. Is the argument supposed to be that medicine *always* has taken adult white men as the standard human? Or are we speaking

of a twentieth-century phenomenon? Or is this tendency deemed to be characteristic of the modern era of the clinical trial—a method of formal experimentation that became prominent only after World War II? Or, perhaps, is the underrepresentation of groups something that emerged only in the 1970s, as an unintended consequence of other reforms that sought to protect "vulnerable populations" from the risks of research? One can find statements by critics of biomedicine over the past two decades that are consistent with each of these conceptions of the historical period under consideration, while many statements are simply too vague to be associated with any specific epoch.

It would be reasonable to imagine that the validity of claims about underrepresentation could be assessed through hard numbers—statistics on precisely who has populated medical studies in the past. I consider numbers of this sort in later chapters, but I defer doing so until then for several reasons. First, even in the period of modern clinical research, the numbers are hard to come by, because researchers were not always careful about collecting or presenting data concerning the social characteristics of their research subjects. (Indeed, one of the benefits of the recent emphasis on inclusion and difference is that federal health agencies are now gathering much more precise aggregate data on research participation—especially in NIH-funded studies, but also in clinical trials leading to drug approvals by the FDA.[6]) As a result of this lack, the available data provide selective and sometimes contradictory views of who was studied, even in the relatively recent past. Second, proponents and opponents of the inclusion-and-difference paradigm have presented competing quantitative analyses of such matters as whether women actually were underrepresented in clinical research in recent decades. Debate on these points has been tightly interwoven into the larger controversy over the desirability of the new policies mandating inclusion. It therefore makes more sense for me to present the numbers in the context of the controversy itself, locating the data within the argumentation of different parties to the debate, so that they may be considered in that context.

Most importantly, when we consider the characteristics of research practices in the past, what is at issue is more than just numbers. The concepts that underlie medical judgments about the use of human subjects—concepts such as sameness, difference, standards, generalizations, vulnerability, ethics, and justice—all have their own complex histories. My goal in this chapter, then, is to excavate some of the meanings of these concepts in order to better situate the analysis of controversy that follows. In doing so, I also mean to develop three points. First, medical researchers have operated with a considerable range of ideas about who constitutes

the ideal, the appropriate, or the acceptable research subject, and these sensibilities also have shifted over time. Second, while therefore it is incorrect to say that white men always have been the "typical and usual" experimental subject, it is true that certain historical developments have encouraged that tendency. And third, given the troubling history of medical conceptions of race and gender, there is a sharp irony associated with the new attention to difference. Recent reformers assume that a medical insistence on difference necessarily advances the interests of historically disadvantaged groups; but the old medical theories of group difference had just the opposite effect, reinforcing oppression and helping to consolidate the very disadvantages we now hope to overturn.

THE HUMAN SUBJECT AS WORKING OBJECT

A key problem confronted by medical experimenters is one that, to a certain degree, faces every practitioner of an experimental science. As analysts in the field of science studies have shown about science generally, the production of trustworthy knowledge out of laboratory settings invariably presumes the creation of standardized "working objects" whose essential characteristics can be claimed to vary little from one laboratory to the next.[7] "No science can do without such standardized working objects," note Lorraine Daston and Peter Galison, "for unrefined natural objects are too quirkily particular to cooperate in generalizations and comparisons."[8] While such working objects are often inanimate, such as scientific instruments and procedures, they may sometimes be living things. For example, Robert Kohler has described how the natural variability of the fruit fly had to be reduced before it could become a reliable experimental object.[9] However, creating a standard human for research purposes is potentially more problematic than standardizing other living things. Whereas scientists could construct a variety of the fruit fly for laboratory purposes, clinical researchers studying humans are obliged to take those humans essentially as they find them. Researchers seeking a standard human "working object" therefore have several options: they may assume that differences between humans are irrelevant for their purposes, they may seek out only those individuals who have chosen characteristics, or they may subject different subpopulations of humans to separate or comparative study.

Of course, researchers may seek to bypass this dilemma—and simultaneously solve the practical problem of finding willing subjects—by substituting experimental animals drawn from other species in place of the standard human. This solution is as old as the history of medical

experimentation, and it continues to be both common and important in many biomedical contexts today. But it, too, may engender controversy over the generalizability of findings.[10] First, it is not always the case that other animals are more easily standardized than humans, and attempts to employ them in medical testing have sometimes run afoul of significant intraspecies variation.[11] Second, there often may be considerable uncertainty about the implications for *Homo sapiens* of findings in other species. For example, a culture like ours today, which has been growing suspicious of extrapolating findings from men to women, adults to children, or white people to people of color may have little patience with the notion that "rats are miniature people."[12] In practice, many medications that appear promising and many findings that appear compelling when tested in animals simply fail to pan out when the experiment is repeated in humans. Thus, most important domains of present-day clinical research, particularly including the clinical trials required by the FDA for the licensing of new drugs, frequently begin with animal testing but ultimately require experimentation in humans for findings to be seen as scientifically adequate and publicly credible.

Let's assume, then, that medical experimentation designed to benefit human beings does require human subjects and that experimenters must, at least in some measure, aim to standardize their practice and the objects of that practice. How have such necessities been made consistent with medical understandings of the human body and human differences?

Medical Theorizing and the Hierarchy of Bodies

In some respects, standardizing the human is a concept foreign to the history of Western medicine. Rather than dealing inevitably in universals, Western medicine has often been preoccupied with differences, both between individuals and between social groups. From a present-day vantage point, this history is both ironic and disturbing. While today's presumption is that medical attention to difference is a beneficial and enlightened response, more typically such attention has both presumed and reinforced a social hierarchy that placed heterosexual European men at the pinnacle. By treating variations between genders and races as something fixed in the body, medical theorists helped to reinforce the perception that social inequalities were a straightforward reflection of the natural order of things.

Western medical theorizing about differences between men and women has a long history. As Londa Schiebinger has noted, the ancient Greek

physician Galen "believed that women are cold and moist while men are warm and dry; men are active, women are indolent."[13] Yet for Galen the female body was essentially a variation on the theme of man: female sex organs were just male organs "turned inward."[14] In eighteenth-century Europe, however, while ideas about male and female sameness did not disappear, strong notions of fundamental medical differences between men and women—of men and women as "opposites"—were used by some medical authorities to breathe new life into claims that women were destined to be socially subordinate to men. As Thomas Lacquer has expressed it, whenever Enlightenment ideas about democracy and equality threatened to erode the old distinctions between men's and women's places in society, "arguments for fundamental sexual differences were shoved into the breach."[15] From a biological and anatomical standpoint, women were often understood to be inalterably different and were portrayed as dissimilar from men in essential and thoroughgoing ways. "The essence of sex," argued the French physician Pierre Roussel in 1775, "is not confined to a single organ but extends, through more or less perceptible nuances, into every part."[16]

These conceptions of female difference did not vanish with the passing of the eighteenth century. In the nineteenth-century United States, physicians seeking to understanding an epidemic of "women's maladies" tended to construe femaleness as almost inherently unhealthy and viewed women as essentially controlled by their reproductive organs. Barbara Ehrenreich and Deirdre English quote one physician who, "addressing a medical society in 1870, observed that it seemed 'as if the Almighty, in creating the female sex, *had taken the uterus and built up a woman around it*.' "[17] Moreover, by the nineteenth century, measurements of European women's skulls and pelvises had led some scientists to conclude that women ranked below men in terms of evolutionary development.[18]

As this history of conceptualizing women's differences suggests, it is important to note that, at least in the nineteenth century, women by no means were "ignored by" or "invisible to" medical practitioners. Carol Weisman has pointed out in her analysis of the history of the women's health movement: "In contrast to the current view that medicine ignores or neglects women, the recruitment of women patients was critical, historically, to physicians' practices, and the development and control of medical treatments for women played a key part in the profession's attempts to establish itself both economically and socially." Indeed, Weisman cites the medical historian W. F. Bynum's claim that, in the nineteenth century, "more often than not, the abstract patient

was referred to as female"—either because women were more likely than men to turn to doctors, or because physicians were more inclined to pathologize women's bodies than men's.[19]

Claims about biological differences also were invoked in nineteenth-century Europe and the United States to justify racial hierarchies in general and the practice of slavery in particular.[20] Samuel Cartwright, chairman of a committee appointed by the Medical Association of Louisiana to report on the "diseases and physical peculiarities of the Negro race," described in a medical journal in 1851 how the skin color of the black man reflected a difference that went all the way inward: "his bile, . . . his blood, . . . the brain and nerves, the chyle and all the humors" were "tinctured with a shade of the pervading darkness." Cartwright argued further that blacks suffered from a deficiency of red blood caused by "defective atmospherization"; but since hard exercise could cure this condition, it followed that slavery improved them "in body, mind and morals."[21] Others argued similarly that rates of insanity were demonstrably higher among freed blacks than among those who remained slaves.[22] Thus, not only did medical beliefs reflect social preoccupations with racial superiority and inferiority, but medicine also played an active role in constructing those very notions of racial character and shoring up the boundaries between races.

Some historians have argued that Cartwright's views may have been somewhat out of the mainstream and that other medical authorities believed that "sickness among blacks and whites differed in terms of degree rather than kind."[23] Still, the arguments of Cartwright and his contemporaries were generally consistent with a broader medical philosophy that was dominant among nineteenth-century U.S. physicians in both the North and the South, which has been called the "principle of specificity." According to this principle, medical therapy "was to be sensitively gauged not to a disease entity but to such distinctive features of the patient as age, gender, ethnicity, socioeconomic position, and moral status, and to attributes of place like climate, topography, and population density." As John Harley Warner has described, "admonitions to heed the various elements encompassed by the principle of specificity permeated therapeutic instruction. . . . Professors routinely taught that [these various] individuating factors . . . modified the character of the disease and the operation of drugs."[24] Thus, notions of racial differences in health and disease were just one dimension of a general insistence on specificity.

As I discuss below, in the late nineteenth century the principle of specificity mostly gave way to notions of medical universalism. Nevertheless, notions of medically inferior races persisted well into the twentieth

century in the United States. A turn-of-the-century work entitled *The Surgical Peculiarities of the Negro* was, according to David McBride, a standard medical reference throughout World War I. In 1910, the *Journal of the American Medical Association* (*JAMA*) published an article by Dr. H. M. Folkes of Mississippi that was entitled "The Negro as a Health Problem."[25] Also writing in *JAMA*, Dr. Seale Harris observed in 1903 that the "[lesser] development of lung tissue and the accessory muscles of respiration among the negroes than for the whites" reflected the fact that "the negro, a century or two ago, was a savage, perhaps a cannibal. . . . [With] the warm, humid atmosphere, less oxygen was required to maintain body temperature, so there was a corresponding lack of development of the lungs of the native African."[26] Allan Brandt has described how physicians pointed to the comparative anatomies of blacks and whites in order to support their claims that emancipation had led to the declining health of the black population.[27] Physicians also were obsessed with the sexuality of black people, who were viewed as "a notoriously syphilis-soaked race," in the words of one early-twentieth-century physician. This presumed susceptibility later provided some of the justification for the infamous Tuskegee Syphilis Study, which was crucially premised on a hypothesis of difference: the study, which denied treatment to black men in rural Alabama in order to track the progression of the disease, sought to investigate whether syphilis might take a different course in blacks than in whites.[28]

Often, white physicians conceived of nonwhite races as posing a special risk of infection to the mainstream U.S. population. Some argued that blacks were, "like the fly, the mosquito, the rats, and mice, an arch-carrier of disease germs to white people."[29] (As Vanessa Gamble has noted, black physicians, by contrast, attributed ill health to poverty and discrimination—not to innate characteristics.[30]) Similarly, in the late nineteenth and early twentieth centuries, public health authorities in cities such as San Francisco conceived of the "alien" Chinese element as a special source of risk to the health of the society and blamed the Chinese for diseases such as syphilis and the bubonic plague.[31] But infectiousness was not the only dimension of racial risk. Under the influence of the eugenics movement, nonwhite racial groups were portrayed as genetically inferior, and intermarriage as a threat to the human gene pool.[32]

Two points are worth emphasizing in considering these troublesome histories of medical conceptions of racial difference. First, while physicians' beliefs reflected the dominant values of the societies in which they lived, their medical arguments were not just passive vehicles for the transmission of racist ideas. Through a specification of difference as rooted in

biology, and by means of the cultural authority invested in the medical and scientific professions, physicians and researchers actively reshaped social understandings of race. For example, in the early twentieth century, when sickle-cell anemia became understood as a disease of "Negro blood," clinicians argued that it was impossible for whites to contract the condition. As both Keith Wailoo and Melbourne Tapper have described, the discovery of apparent cases of sickle-cell anemia in whites led to classificatory dilemmas that were resolved by inferring the existence of some hidden African ancestor in the "white" individual's family tree. In this way, medical theorizing helped not only to shore up the racial divide, but also to give new meaning to what it meant to be black: part of the very definition of black identity was the susceptibility of "Negro blood" to illnesses such as sickle-cell anemia.[33]

The second point is that invidious notions of biomedical racial differences persisted in some quarters in the United States past the formation of a scientific consensus, in the interwar years, that notions of racial superiority and inferiority lacked any scientific basis.[34] Indeed, even the association of Nazism with eugenics and racialized medical science failed to dispel completely the notion that racial hierarchies were "in the blood."[35] In 1942 a committee of physical anthropologists complained that the maintenance of segregated blood banks under Jim Crow laws was embarrassingly reminiscent of practices "based on the Nazi theory of race."[36] Blood plasma was also kept segregated in the U.S. military throughout World War II, though apparently less because military physicians perceived a medical necessity to do so than because some members of Congress thought that "the argument that the blood of whites and blacks was interchangeable [was] a Communist plot to 'mongrelize America.' "[37]

The residue of nineteenth-century biological conceptions of race is also evident in the antiquated terminology that was used in the United States until 2004 in the indexing codes for medical journal articles. The National Library of Medicine's "Medical Subject Headings" included the indexing term "Racial Stocks," which was subdivided into "Australoid Race," "Caucasoid Race," "Mongoloid Race," and "Negroid Race." A search of the medical literature reveals that from 1990 through 1999, nearly 13,000 medical journal articles about human populations were coded with the indexing term "Racial Stocks," and thousands bore the more specific codes. Beginning with articles published in 2004, these subject headings were eliminated and replaced with "Continental Population Groups."[38]

Experimentation and Difference: Sex, Age, Class, and Race

The preceding section provides a schematic overview of the significance of certain kinds of difference for medical theory and practice in the eighteenth and nineteenth centuries. But it mostly leaves to one side a consideration of the more specific issue of medical experimentation. In fact, during this period, many physicians were active experimenters, eager to test new substances for their therapeutic properties. As David Rothman has noted, "the idea of judging the usefulness of a particular medication by actual results goes back to a school of Greek and Roman empiricists," but it was in the eighteenth century that human experimentation began to make "its first significant impact on medical knowledge . . . , primarily through the work of the English physician Edward Jenner."[39] On whom did doctors perform such experiments? Although it was common to try out such substances first on animals, physicians considered it necessary to proceed to experiments with humans. In many cases, physicians then experimented initially on themselves.[40] But medical patients often found themselves serving as experimental subjects whether that was their choice or not.

In a recent article, Schiebinger provides an illuminating look at how understandings of difference and conceptions of the natural body affected experimenters' selection of subjects in eighteenth-century Europe. She notes: "To some extent, the choice of subjects was simply arbitrary. As with dissection, physicians and surgeons used any bodies they could lay their hands on (perhaps legally and morally, perhaps not)." Prisoners, hospital patients, orphans, and soldiers were among those most likely to be experimented upon, and "from a medical point of view, there was nothing special about these bodies, except their availability." At the same time, conceptions of sameness and difference were sometimes important to experimental practice. Schiebinger presents evidence that many physicians routinely recorded the age and sex of those experimented upon, reflecting beliefs that therapies might have different effects on children and adults and on women and men. But despite the extraordinary importance attached to social class distinctions in the eighteenth century—and the extent to which class differences were seen as inherited—experimenters in that case "assumed an interchangeability of bodies among Europeans" and were content to trust that experiments conducted on the poor held medical relevance for the rich. Neither was race considered a barrier to extrapolation throughout most of the eighteenth century. By the end of the century, however, as strong biological notions of race became more

widespread, experimenters began to worry whether the bodies of Africans were representative of humankind. Such worries did not, however, prevent physicians from continuing to experiment on slaves, whose "availability" for such purposes was simply too appealing to be ignored. Thus, in the eighteenth century, certain differences were seen as barriers to generalization while others were not (and these distinctions themselves varied somewhat over time), but practical exigencies concerning the supply of experimental bodies often overrode all other considerations.[41]

As the work of other historians has suggested, by the nineteenth century, conceptions of racial difference had become so entrenched in medical practice that those who wanted to experiment on slaves—and many did, once again for the simple reason of their availability[42]—were obliged to make arguments about human similitude that challenged the conventional wisdom. A case in point is the surgical experimentation conducted by Dr. J. Marion Sims, the nineteenth-century, South Carolina—born "father of gynecology." Sims developed an important and revolutionary surgical procedure—a remedy for vesico-vaginal fistula—through experimentation on slave women. These women were provided to him by their owners, though in at least one case Sims purchased a slave specifically in order to experiment upon her. That Sims's choice of subjects was dictated by expediency and the social organization of power in the antebellum South is obvious enough. Yet it was crucial to the larger success of Sims's work that he be able to argue that the procedures that he elaborated on the bodies of black women would be equally efficacious when applied to white female patients. From an intellectual standpoint, Sims thus found himself in direct opposition to contemporaries such as Cartwright who emphasized the "peculiarities" of the Negro race. Instead, as Deborah McGregor has noted, "Sims assumed that female anatomy was homologous between whites and blacks" and that extrapolations could therefore be made from the latter to the former.[43]

A similar example would be the use, in the antebellum South, of black corpses for purposes of anatomical dissection. At a time when medical institutions resorted to contracting with grave-robbers to obtain corpses, black bodies were simply more vulnerable to expropriation. "In Baltimore the bodies of coloured people exclusively are taken for dissection," a visitor to the United States from France commented in 1835, "because the Whites do not like it, and the coloured people cannot resist."[44] But the interesting outcome is that nineteenth-century physicians developed ideas about human anatomy based largely on the study of African Americans. As Robert Blakely and Judith Harrington have observed: "It is one of the ironies of medical history that, although blacks were generally

regarded as 'inferior' or even 'subhuman,' their corpses were considered 'good enough' to use in the instruction of human anatomy."[45] Here again, the "availability" of black people for medical purposes resulted in the fact that, in specific contexts, they, rather than whites, effectively served as the standard human.

There are also plenty of more recent instances of medical experimentation on racial and ethnic minorities that have resulted in the extrapolation of findings from people of color to whites: for example, the use of women in Puerto Rico for some of the early studies of the Pill in the 1950s and the present-day interest of some geneticists in studying African Americans because they possess some of the genetic mutations that are the "oldest" and hence the most universally distributed.[46] Who, then, is the "standard human"? These latter cases seem to contradict the claims of those who say that only the most socially privileged groups have served as the standard and that others either have been invisible or have been studied only in terms of their difference. These examples demand attention to those moments when researchers not only have (for whatever combination of sensible and dubious reasons) placed women and minorities front and center, but also have been willing to say that results thereby obtained will apply to humans generally.

Captive Populations, Abuses, and the Rise of Protectionism

The use of specific groups as "captive populations" for research purposes merits further consideration. While it is rightly assumed that captive populations typically come from socially disadvantaged or exploited groups, this is not always the case. Occasionally, availability has meant the study of relatively socially advantaged groups. In an article aptly titled "Using the Student Body," Heather Munro Prescott has described the ubiquitous reliance on college students as samples of convenience for academic researchers. As Prescott notes, "the assumption that undergraduates are natural research subjects is so deeply embedded in both the history of and present-day thinking on human experimentation that it is difficult to separate discussion of student subjects from that of other healthy volunteers." Studies of students have been used to establish a range of "baseline" physiological measures and standards of normality, including "the normal ranges for blood pressure, lung capacity, pulse rate, basal metabolism, and other physiological processes." For these purposes, students have been considered "ideal" in the double sense: a handy "captive population," but also ideal specimens of human normality. In this case, "captive" does not necessarily mean "vulnerable"; as Prescott makes clear

in considering patterns of research at schools such as Harvard in the first half of the twentieth century, student research subjects may have treated their professors with deference, but they were not their social inferiors, and they did have a capacity to look out for their own interests. And researchers interested in making claims about "normality" often preferred students from privileged backgrounds, who were perceived to be "the best representatives of normality by virtue of their race, class, and gender."[47]

As interesting as this case of "the student body" may be, it should be emphasized that the captive populations that have served as the basis for medical generalization typically have consisted not of the socially privileged but of the relatively disadvantaged. Institutionalized populations of various kinds—soldiers, incarcerated prisoners, the mentally ill, retarded children—along with the poor in general, continued to provide much of the human raw material for medical studies at least through the mid-1960s.[48] After World War II, with a huge influx of funds to the NIH, the scope of medical experimentation in the United States increased enormously. But from 1945 to 1965—a period that David Rothman has called "the Gilded Age of research" in ironic recognition of the laissez-faire attitudes that prevailed—this expansion sparked little reflection on the rights of research subjects.[49] Indeed, although the Nuremberg trials after World War II had provided graphic evidence of the horrific uses to which medical experimentation could be put, only in the 1960s, with the publication of reports of widespread abuses of patients in high-profile, U.S. medical experiments, did many policy makers begin to assert that stricter measures were needed to safeguard human subjects in the United States.

Rothman has chronicled the impact of the appearance in the *New England Journal of Medicine* in 1966 of a whistle-blowing article by a doctor named Henry Beecher, who compiled a list of twenty-two examples of studies published since World War II that struck him as patently unethical: withholding penicillin from soldiers with streptococcal infections in order to study alternative means of preventing complications; feeding hepatitis virus to residents of a state institution for the retarded; injecting live cancer cells into elderly and senile hospital patients without telling them the cells were cancerous; and so on. As Beecher emphasized, these examples did not come from the fringes of medicine; they were conducted by well-established researchers at prominent institutions, and the results had been written up in all the best journals.[50]

Six years later, in 1972, Associated Press reporter Jean Heller broke the story of the U.S. Public Health Service study of "Untreated Syphilis in the Negro Male"—commonly known as the "Tuskegee Syphilis Study" because government researchers collaborated with physicians at the Tuske-

gee Institute in Alabama. A 1929 survey by the U.S. Public Health Service had found an exceptionally high prevalence of syphilis in Macon County, Alabama—a rural county with high rates of poverty and low rates of education among its heavily black population. Therapies consisting of mercury and arsenic compounds, though dangerous, were known to be of at least limited benefit to patients with syphilis, but these medications, like many others, did not typically find their way to rural Alabama. Therefore, as James Jones has described the rationale, because most of those with syphilis in Macon County went untreated anyway, it seemed sensible to researchers to observe the consequences.

After recruiting 399 men with syphilis and 201 controls, investigators tracked the progression of the disease over four decades, minutely recording its devastation in the form of skin ulcers, deterioration of the bone structure, problems in motor coordination, blindness, and death. Incredibly, even after the discovery in the 1940s that syphilis could be treated effectively with penicillin, participants were never given antibiotics for their condition, and they were actively discouraged from seeking medical attention from other doctors, lest they inadvertently gain access to such drugs. In the end, somewhere between twenty-eight and one hundred of the men died as a direct result of syphilis or its complications. To complete the picture of abuse, participants mostly were unaware even that they were subjects in an experiment; they were led to believe that they were receiving treatment for their "bad blood." Invasive tests such as spinal taps, designed only to gather data for the study, were presented to research subjects as "special" treatments intended for their medical benefit.[51]

The story of the Tuskegee study is crucial to the arguments in this book for several reasons, two of which need to be emphasized here.[52] First, as mentioned previously, Tuskegee was justified in part through the logic of racial difference. As untreated syphilis had already been studied in a white population in Norway, researchers claimed that part of the goal of the study was to determine whether the progressive course of syphilis in black people was similar to, or different from, its known trajectory in whites.[53] Thus, from the vantage point of the present era, in which it is frequently claimed that racial minorities will benefit from research that does not presume that whites and people of color are medically equivalent, the episode serves as yet another troublesome reminder that medical research premised on racial differences can sometimes serve stigmatizing and dangerous ends.[54]

A second aspect of the significance of the Tuskegee Study is that the reaction to it in the 1970s was emblematic of a shift in "common sense"

on the part of medical professionals and the broader society—a transformation that held important implications for the study of "vulnerable populations." The Tuskegee study was in some respects an extreme case, but, as Beecher's article had made evident (and as more recent revelations, ranging from radiation experiments to novel treatments of the mentally ill, also have underscored[55]), it was by no means unique in the annals of experimentation in the twentieth-century United States. Moreover, the study had never been a secret. Although news of the study was greeted with horror and disbelief when Heller "exposed" it, reports from Tuskegee had appeared in the pages of medical journals on a regular basis since 1936. As late as 1969, a CDC committee had reviewed the study and determined that it should be permitted to continue. Therefore, Tuskegee was not a "dark secret" that was suddenly brought to light. Rather, the publicizing of Tuskegee marked, and further propelled, a changed understanding of legitimate practice with regard to experimentation on the socially disadvantaged—a new moment in what Sydney Halpern has described as the ever-evolving formation of researchers' "indigenous moralities," as well as a new phase in public debate.[56]

In the 1970s in the United States, the public attention drawn to the ethics of human experimentation swelled into a wave of governmental reform. This wave crested with the enactment of formal, legal protection of the rights of experimental subjects, along with a new conception of participation in research as a burden which, therefore, must be distributed as equitably as possible in society.[57] With the passage of the National Research Act of 1974, researchers became obliged to comply with procedures established by the NIH's new Office for Protection from Research Risks, to submit their protocols beforehand to local "institutional review boards" (IRBs) that would ensure that human subjects were not placed at undue risk, and to document the process of obtaining informed consent from their subjects. The new thinking was further enunciated in the "Belmont Report" published by the National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research in 1979, which outlined the ethical principles that ought to guide medical research with human subjects: justice, respect for persons, and beneficence.[58]

A distinguishing feature of this regime of regulation was its emphasis on the protection from harm at the hands of the research enterprise of what were now officially defined as "vulnerable populations," including children, prisoners, the poor, and the mentally infirm.[59] In the wake of severe birth defects in the children of women (mostly in Europe) who received the drug thalidomide during pregnancy, U.S. regulators increasingly came to view women, or at least their potential fetuses, as yet

another vulnerable population meriting protection. In 1977 the FDA instituted a rule formally excluding women "of childbearing potential" from many drug trials, out of concern that an experimental drug might bring harm to a fetus if a woman became pregnant during the course of a clinical trial.[60]

It is easy to see how this new regulatory climate, designed to protect against abuse of subjects, may have resulted in the reduced representation of women, children, and racial and ethnic minorities in biomedical research. But it is important to observe how relatively recent these reforms were at the time that complaints about underrepresentation became rampant in the late 1980s. Like many reforms, they may inadvertently have created a new set of problems to which a subsequent generation of reformers then found themselves obliged to respond. But it would be incorrect to attribute any long-term tendencies in medical knowledge-making practices (in terms of who gets studied or who serves as the standard) to procedures that were introduced only in the 1970s.

FROM THE PRINCIPLE OF SPECIFICITY TO *L'HOMME MOYEN*

In attempting to make sense of the complexity of historical beliefs and practices with regard to medical theory and experimentation, this chapter has examined several interlocking historical trajectories: the evolution of medical conceptions of difference, ideas about drawing generalizations from experiments on human, the exploitation of captive populations, and the rise of modern notions of protecting human subjects. One additional but equally crucial history remains to be discussed: the entry of modern statistics into medicine and, more broadly, the role of quantification, measurement, and standardization in promoting notions of medicine as a science.

When the French mathematician Adolphe Quetelet announced the new science of "social physics" in 1831, its central concept was what he called *l'homme moyen*, the average man. This man would have not just an average height, weight, education, and length of life, but also an average propensity to marry, commit suicide, or engage in criminal acts. As Gerd Gigerenzer and coauthors have noted, Quetelet fully understood that *l'homme moyen* was an abstraction who existed nowhere in reality: "But abstraction was essential to social science. Real individuals were too numerous and diverse for psychological study to contribute much to an understanding of the social condition."[61]

Under the logic of the "principle of specificity" that I described as prevailing in U.S. medical practice much of the nineteenth century, sta-

tistical constructs of this sort would have found little place. As proper medical treatment was presumed to be different for men and women, Northerners and Southerners, the rich and the poor, and so on, few physicians would have been comfortable diagnosing or prescribing for *l'homme moyen*. However, in the latter part of the nineteenth century, with the increasing adoption of European theories of scientific medicine, U.S. physicians gradually abandoned the notion that treatments should be tailored to the idiosyncratic constitutions of patients, in favor of the idea that each specific kind of illness required a distinctive treatment that might be applied universally to sufferers.[62] Increasingly, medical practitioners began to adopt conceptions of the human individual that were influenced by the rise of the sciences of statistics and probability.

The move away from the principle of specificity and the importation of statistical conceptions of humanity into medicine presaged a broader tendency. Over the course of the twentieth century, a host of developments converged to encourage a thoroughgoing, though often resisted, standardization of medical practice: the rise of modern methods of pharmaceutical drug testing and drug regulation, the development of epidemiological studies based on notions of statistical risk, the codification of international classification systems for morbidity and mortality, the increased reliance on standard protocols and expert systems, the rise of evidence-based medicine as a social movement within biomedicine, and the growth of managed care as a system of rationing and surveillance. All these developments have privileged a conception of medicine as "science" (consistent, predictable, and transparent) over a competing conception of medicine as "art" (dependent on intuition, experience, and embodied skills, and respectful of the particularities of individual patients).[63] By no means have conceptions of artful medicine disappeared, and physicians may often invoke the idea when defending their professional autonomy against bureaucrats or insurance companies. But the significance of standardization to present-day medical practice would be hard to deny.

The standardization of medical practice has often carried with it a strong presumption that the object of medical attention—the patient—could likewise be conceived of in relatively standard and universal terms. In spite of resistance, such views have informed medical education and training in various respects. It is telling that the actors who are trained to simulate disease symptoms for the benefit of students in medical school classrooms are referred to as "standardized patients." Now used in medical schools across the United States to provide students with "real-life" experience, standardized patients are labeled as such because the actors endeavor to present a consistent simulation each time they perform.[64]

Mechanical simulators of patients—such as the virtual reality surgical simulators currently used in some medical teaching settings—carry this standardization to an even greater degree, helping thereby to standardize the practices being taught by means of the simulators.[65]

Importantly, such standardization has implications for the key question of which sort of person actually proves to be the focus of biomedical attention. In particular, the standardization of the patient in medical training has sometimes coincided with a privileging of the male body over the female, precisely as recent critiques of biomedicine would suggest. For example, in a study of the classic medical reference book *Gray's Anatomy* from 1858 to 1998, Alan Petersen found remarkable consistency, over 140 years, in the manner by which "the male body has been . . . posited as the standard or norm in both illustrations and textual descriptions." Petersen's point was not, however, that the female body was absent or invisible in *Gray's Anatomy*. Rather, Petersen found that the volume appeared to call attention to women's bodies precisely insofar as they differed from men's. These comparisons emphasized "the superiority of the male body, which is seen as stronger, more fully developed and more active than the female body."[66] Similarly, in her analysis of the mechanical simulators used for teaching purposes, Ericka Johnson found that "the male body is used as the norm and the female body represented only when it differs from the male, and then only in the 'parts' which are 'importantly' different."[67]

In considering the trend toward standardization in medicine, one might well object that the modern notion of the "risk factor"—itself also a product of the statistical revolution in medicine and epidemiology—complicates any simple story of the standard patient. After all, it has become increasingly well understood that each patient may have a distinctive risk profile that makes him or her more or less susceptible to particular diseases.[68] Yet the development of risk guidelines in practice often has reflected a tendency to universalize in potentially inappropriate ways. A classic case is that of the famous Framingham Study, an observational study of heart disease conducted in a small community near Boston that began in 1948 and has continued for more than half a century. Framingham data have had an extraordinary impact on medical practice, and they are the basis of the standard risk assessment tools for gauging the impact of high blood pressure, blood lipids, smoking, physical activity, and obesity on the risk of developing heart disease.[69] But Framingham was hardly America, even in 1948 when the research began. As one of its directors observed in 1980: "There were virtually no blacks or Orientals, and the composition of the white population was not necessarily that of white populations elsewhere."[70] Moreover, because participants were

aged 36–68, investigators missed many cases of heart disease in women, who tend to develop it later in life.[71] Thus, even tools designed to differentiate within populations (by distinguishing those at higher and lower levels of risk) may inappropriately homogenize the population to which they are applied, if developed on the basis of unrepresentative samples.[72]

The Randomized Clinical Trial and the Problem of Variation

A crucial stage in the standardization of the patient, the quantification of medical research, and the increased reliance on the human subject as an experimental object was the emergence of the randomized, controlled clinical trial as a distinctive kind of medical experiment. Formally adopted by medical authorities after World War II, and subsequently made part of the legal process of regulatory decision making about the safety and efficacy of drugs, the methodology of the randomized clinical trial has sought to place clinical practice on a solidly scientific footing; it is often called the "gold standard" for establishing the effect of any medical treatment on humans, and it is considered more reliable than observational methods.[73] By comparing results in two (or more) groups of patients—divided typically into a "treatment arm" and a "control arm"—investigators could determine with greater confidence whether the apparent outcomes of a medical intervention were genuinely due to that intervention rather than to some other cause or to chance. Randomization—the random assignment of each study participant to either the treatment arm or the control arm of the study—decreased the possibility of investigator bias in placing patients into groups, while also making possible the use of statistical tests to assess the significance of results. Randomization also was the prerequisite of the successful use of another important technique to avoid bias: "double blinding," in which neither the investigator nor the subject knows which subjects have been assigned to which arm of the study.[74]

How did the development of the randomized clinical trial affect who served as the experimental subject of choice in medical research? In fact, this innovation in medical knowledge-making held a variety of consequences for the demographics of human experimentation. As clinical research became more scientifically grounded and more central to the image of modern medicine, and as more and more funding poured into clinical trials from both the government and pharmaceutical companies, it is reasonable to imagine that researchers wanted to use the new techniques to address problems affecting mainstream or socially privileged groups, such as a perceived epidemic of heart disease in men. Similarly, it

is reasonable to suppose that drug manufacturers wanted to study those groups in society with the greatest ability to afford their remedies. At the same time, the methodological advances themselves obviated some of the need to worry about precisely who was in one's study. Specifically, the technique of randomization provided at least a partial solution to the problem at the heart of the discussion in this chapter: how to deal with the fact that patients vary. In theory, at least in trials with large numbers of subjects, any factor that might, unbeknownst to investigators, cause certain patients to respond differently—say, a patient's age or previous medical history—would end up being more or less equally represented in both arms of the trial after randomization and therefore would not affect the trial's overall results. Thus, randomization could solve some of the difficulties associated with human variability without requiring any additional efforts on the part of researchers to recruit particular kinds of individuals.[75]

As researchers and statisticians refined the rules governing randomized trials, they developed additional procedures with implications for who ends up populating a study. Importantly, the protocols for all trials came to specify "inclusion/exclusion criteria"—formal rules designating eligibility for participation in a trial. All trial protocols must state some such criteria, but the question is just how specific they need to be in any given case. Strict inclusion/exclusion criteria are sometimes used to create a more standardized and homogeneous research population for a study, on the argument that the more researchers reduce the number of variables that might affect a study, the easier it will be for them to distinguish "signal" from "noise." On the other hand, experts on clinical trials who are concerned with the "external validity" of a trial—the significance of its findings for large numbers of people in everyday life settings, and not just for the smaller number of people who happened to participate in the clinical trial—often tend to favor more relaxed entry criteria. These experts argue that heterogeneous research populations are better models of actual conditions and that studies with homogeneous populations may result in scientifically elegant experiments that lack much real-world significance.[76] There is no solution, in principle, to the problem of how to resolve this trade-off between experimental rigor and generalizability of findings. As late as 1983, one authority on clinical trials, Dr. Alvan Feinstein, writing in the *Annals of Internal Medicine*, described an ongoing war between these two conceptions of clinical trials, which he labeled "fastidious" and "pragmatic," respectively.[77] Others have used the terms *efficacy* and *effectiveness* to distinguish between two different, potential goals of clinical trials—on one hand, establishing a statistically

significant finding through experimental procedures; on the other hand, demonstrating the worth of medical interventions on a large scale, outside the experimental situation.[78]

Once again, it is worth considering the implications of these innovations in technique for the issue of who has been most or least likely to be represented in biomedical research populations. In the next chapter, I return to the question of homogeneity versus heterogeneity in subject populations and consider the implications for debates about the underrepresentation of racial and ethnic minorities, the elderly, children, and women in medical research. For the moment, it is enough to observe that, to the extent that researchers adopt the goal of recruiting a homogenous subject population, the possibility clearly exists that some social groups will find themselves excluded from studies.

Complexities

This chapter has emphasized a consistent medical preoccupation over several centuries with a set of problems: how to assess the medical relevance of differences between groups; how and whether to make generalizations about human health and illness; and how to decide which human beings are appropriate, or preferred, or deserving, or exploitable as experimental objects. It is important to observe the continuities in these preoccupations. But it is equally crucial to chart the impact of historical changes and the continued evolution of various relevant notions—of what a proper medical experiment looks like, of what we imagine group differences to signify biologically, of what rights we consider human beings intrinsically to possess. Shifts in such sensibilities, combined with scientific and technological advances in medical and statistical capacities, have held important implications for experimentation on humans—or particular subpopulations.

Let us take stock of the arguments developed in this discussion. First, though one can certainly find examples reaching into the past of cases in which medical professionals and researchers have presumed a standard human, significant standardization of the human for medical purposes is mostly a development of the twentieth century. Indeed, much of the history of Western medicine is a history of difference-making, importantly including gender difference and, by the late eighteenth century, racial difference. In the nineteenth-century United States, doctors by no means ignored women or people of color, nor were such groups medically invisible—to the contrary, physicians and researchers were often preoccupied with making claims about them. However, while in present-day

discourse a medical attention to difference is considered progressive and liberatory, the past history more typically suggests a much more deeply problematic tendency—the attempt to inscribe notions of superiority and inferiority, and normality and pathology, on the bodies of different races and sexes.

Second, practical issues governing the availability of subjects for medical experiments often have overridden medical concerns about extrapolation: when the only subjects available were those from captive or vulnerable populations, then experimenters many times have been willing to "take what they could get" and, moreover, to maintain that the information thereby gleaned is of some general relevance. However, recent concerns with the rights of human subjects have inspired moves to protect "vulnerable populations" from the risks of medical experimentation; and while such protection may have lessened the tendency to rely on "available" subjects, it also may have resulted in reduced representation of women, children, the elderly, and racial and ethnic minorities (among others) as subjects in clinical research.

Third, with the importation of statistics into medicine, standardized notions of the medical patient and the research subject gradually have taken hold within medicine, especially in the twentieth century—potentially promoting a bias in favor of seeing white male adults as the standard type of human. But this standardization has proceeded unevenly and in the face of resistance and contrary tendencies. Also, some techniques associated with the modern clinical trial, such as randomization, appear to bypass the problem of confronting variation among research subjects. However, at least some investigators have sought to use strict inclusion/exclusion criteria as a way of ensuring a more homogeneous study population; and this also has the potential to contribute to the underrepresentation of particular groups as research subjects.

In short, this chapter has identified several paths by which various groups, such as women and racial and ethnic minorities, have sometimes come to be underrepresented as research subjects. But at the same time, the notion that the heterosexual, middle-aged, white male has served as the biomedical standard and that all other groups have been essentially invisible, is simply too sweeping when posed in universal terms. Claims such as Healy's—about "the orthodoxy of sameness and the orthodoxy of the mean," about treating "the average American male" as the "normative standard" and extrapolating from him to others—need to be qualified and placed in context.

The more, it seems, that we examine diverse instances of medical experimentation in the past, the less confident we may feel about making

general claims about who has been included, who has served as a standard, and when the extrapolation of findings across groups has been deemed reasonable or advisable. By reviewing this history and cataloguing various examples, my point is not to doubt the observation that the adult white male frequently has been taken as the standard human type in medical thinking and practice. But the broad-brush assertion that the adult white male had become the universal human subject fails to do justice to the varying particularities of research designs, the competing ideas about human sameness and difference, and the creative and sometimes troubling ways in which researchers have responded to practical exigencies in order to carry out their work. Nor has this whole past history of contingency and particularity been unearthed by medical historians in a systematic or comprehensive way. A more adequate understanding of the diverse imaginings of the human subject awaits a good deal more scholarly investigation.[79]

This historical review sets the stage for a consideration, in the next chapter, of the rise in the 1980s of "antistandardization resistance movements" that demanded greater inclusion of underrepresented groups in medical research. First and foremost, the preceding discussion is meant to provide helpful context for thinking about this movement's claims with respect to medical standardization, representation, and generalization. But in addition, many of the examples presented here provide an important and ironic counterpoint to the argumentation of reformers: In the past, claims about differences between men and women and between whites and other racial groups were used to bolster conceptions of the innate superiority of white men. By contrast, the reformers of recent decades have sought to use evidence of biological differences precisely as a grounding for antisexist and antiracist political activism.