Katherine M. Flegal

ANNUAL REVIEWS

Annual Review of Nutrition

A Female Career in Research

Katherine M. Flegal

Stanford Prevention Research Center, Department of Medicine, Stanford University School of Medicine, Stanford, California, USA; email: katherine.flegal@gmail.com

Annu. Rev. Nutr. 2022. 42:1–19

First published as a Review in Advance on April 1, 2022

The *Annual Review of Nutrition* is online at nutr.annualreviews.org

https://doi.org/10.1146/annurev-nutr-062220-103411

Keywords

NHANES, Matilda effect, unconscious bias, public health, statistics, anthropometry, autobiography

Abstract

After a long career at the National Center for Health Statistics, I retired and joined the Stanford Prevention Research Center as an unpaid associate. I was once described by a former US Food and Drug Administration commissioner as "one of the great epidemiologists." The chair of the Harvard nutrition department, speaking on National Public Radio, once described my research as "rubbish." Both may be exaggerations. Here I address some of the events that led to these contrasting descriptions. I also address the extent to which the so-called Matilda effect may have influenced my career. Are women in science on an equal footing with men? The Matilda effect suggests not. Unlike the Matthew effect for scientists, whereby those of higher prestige accrue a disproportionate share of recognition and rewards, the Matilda effect proposes that women scientists are systematically undervalued and underrecognized. I could never get a faculty job and was often treated like an underling. Nonetheless I persevered to publish highly cited research on several high-profile and sometimes controversial topics. Though overt sexism in science and workplaces has diminished over the course of my career, progress toward eliminating unconscious bias has been slower. The Matthew and Matilda effects are still powerful forces that distort incentives and rewards in science.

Contents

INTRODUCTION: SEEING AND BEING SEEN

The 1986 holiday party at the University of Michigan (UM)'s biostatistics department, where I had a research appointment, was a small affair at someone's house. My spouse, a musicologist with no connection to the department, who looked reasonably professional and professorial and a little too old to be a grad student, had accompanied me. Crossing the room, I noticed one of the biostatistics professors looking at him and asking the department chair, "Who's that guy?"

> "That's Katherine Flegal's husband," the chair replied.
>
> "Who's Katherine Flegal?"

For more than a year, I had walked by this professor's office nearly every day in this small department, apparently unseen. All my husband had to do to be noticed was to stand around at a party. But I—an actual member of the department—was invisible.

The #MeToo movement focused much-needed attention on the alarming frequency of sexual harassment and sexual assault—and on the power structures that have prevented those responsible from being held accountable. Those headline-grabbing cases are at one end of a spectrum of discrimination that encompasses the far more common problems of unconscious racial and gender bias and of systemic racism and sexism: biases that are baked into long-held traditions, practices, and rules (64).

When I was invited to write this reflection on my career in science, I began thinking about the narrative arcs that success stories such as these often follow: An enterprising researcher gains status as their good work is recognized; they become a professor, dean, or director; they build up a coterie of trusted colleagues and bright fellows and grad students; they win awards; they are appointed to high-profile committees; and they become members of prestigious organizations.

That is not my story. For me, there was no clear arc—no steady accumulation of accolades, support, and security. I might say I took the road less traveled, but really it was more like driving down a country road with detour signs, roadblocks, and speed traps.

Looking back at this road, I started thinking about the Matthew effect and then the Matilda effect. The Matthew effect, described by sociologist Robert Merton in 1968 (55), refers to the way that those with initial advantages receive disproportionally more additional advantages of fame and status, sometimes summarized as the rich get richer and the poor get poorer. Rather than the Matthew effect, the dominant phenomenon that I and many other women in science have experienced is the lesser-known Matilda effect, described decades ago by Margaret Rossiter (69). Unlike the Matthew effect, whereby those of higher prestige gradually accrue more recognition and rewards, the Matilda effect proposes that women scientists are systematically undervalued and underrecognized. Women are underrepresented in scientific awards (49), citations (8), and various forms of support that lead to high research productivity (1).

A researcher who changed genders during his scientific career observed the effects of unconscious bias against women firsthand. Stanford neurology professor Ben Barres earned his MD, PhD, and first faculty appointment under his birth name, Barbara. At age 42, he transitioned from Barbara to Ben and from a female to a male identity. Before transitioning, Barres had assumed that he had been treated like everyone else, but afterward he saw how profoundly his life had been shaped by his gender and how much his ideas, contributions, and authority had all been devalued when he was perceived as female (4, 5). "By far the main difference I have noticed," Barres wrote in *Nature* in 2006, "is that people who don't know I am transgendered treat me with much more respect" as a man than they did when he was a woman. "I can even complete a whole sentence without being interrupted by a man" (4, p. 135). In his autobiography, Barres cites a counselor working with transgender people. The counselor said that her most difficult challenge was to help men who transitioned to women understand that the perceptible drop in social status they experience after transitioning happens not because they are transsexual, but because they now present as women (5, p. 100).

The absence of a classic arc makes my professional story challenging to tell—but it also presents an opportunity, because I believe my story is one in which many women working in science will find something familiar, and one that some men in science may find eye-opening.

For me, as for many women scientists of my generation and indeed many today, unconscious biases against women too often made us and our contributions barely visible. Despite these many challenges, I was able to contribute advances in research methods and robust findings that influenced and sometimes surprised the research community. I take pride in the large number of citations of my work, the subsequent analyses of public health data my work helped provoke, and the conversations with individual scientists who tell me how my papers influenced their research. I also experienced some of the constraints—but also the odd advantages—that come with work at a government agency, particularly one that saw research as a luxury rather than as its core mission.

FINDING SCIENCE: A USEFUL DETOUR THROUGH DATA AND CODE

I was born and raised in Berkeley, California, a small town with a large university, the University of California (UC). My home life was unusual at the time in our middle-class neighborhood; my mother, my brother, and I, as well as my aunt and my cousin, all lived with my maternal grandparents. I was the only child in my elementary school class with divorced parents; my father was absent from my life. My family members on both sides were well educated and it was taken for granted that I would go to university, but with the assumption that it would not lead to much. My mother, with a college degree, worked as a secretary; my husband's mother, also a college graduate, worked as a secretary, as did the college-educated mothers of several of my friends. Being a secretary was one of the few career paths open to middle-class women, along with being a teacher or a nurse. Sex discrimination in employment and academia was perfectly legal and widespread. Job ads in newspapers were divided into "Help Wanted—Men" and "Help Wanted—Women."

I attended UC and had a student job in the library as an undergraduate. There were also higher paid student-supervisor jobs, but only male students were allowed to be supervisors. No one objected to this policy or felt it was inappropriate or unusual. I was interested in archaeology as an undergraduate anthropology major. Then, as now, there were numerous summer field archaeology schools one could attend to learn the essential skills of field work. With one exception, that of a school run by a woman archaeologist, all of these field schools admitted only male students and did not accept women. Again, this was viewed not as discrimination but simply as a fact of life. The accepted wisdom among female archaeology students was that there was just one sure way for a woman to become an archaeologist: marry one.

It was 1962: Women accounted for 10% of doctorates, 5% of medical school graduates, and just 3% of law school graduates. One spokesman for a medical school in that era said (11, p. 20), "Yes indeed, we do take women, and we do not want the one woman we take to be lonesome, so we take two per class." Although it was not impossible for women to get into medical school or law school, their chances of getting admitted were very low.

I graduated from college in the spring of 1967 with a bachelor's degree in anthropology, amid considerable chaos on the UC campus. I left immediately for a small town in Turkey where I lived for seven months, staying with a family and tutoring English. Upon my return, I took a job fact-checking at *Ramparts*, a left-wing magazine, and after being laid off, I found a temporary job as a clerk in the county assessor's office. The clerks, almost all female, were supposed to rotate making coffee in the break room; the two male clerks successfully petitioned to be excused from performing this humiliatingly feminine task.

It was still the case that few employment opportunities were open to women. I didn't want to be a secretary, nurse, or teacher, and I began thinking seriously about how to get some kind of viable job. I decided that I should try to become what was then called a computer programmer. This was after the time when coding was considered a female secretarial activity but before the time when programming became almost 80% male (74, 75). In the 1960s, businesses and government agencies had started to transfer major parts of their operations to huge mainframe computers. They needed programmers, but computer science was barely addressed in universities and then only as part of electrical engineering or applied mathematics. One of the few ways to learn programming was to get hired as a trainee on the basis of aptitude tests.

I was good at aptitude tests and eventually got a position as a programmer trainee at the Alameda County Data Processing Center. I was one of two women and three men hired as programmers for the county's three IBM System/360 mainframes, huge machines weighing more than 7,000 pounds apiece. Our six months of intensive training involved lectures and reading and was extremely thorough. We learned excellent programming skills and acquired a good understanding of the computing environment and complex job control language before graduating to become full-fledged programmer analysts. The work environment was remarkably gender neutral, with roughly equal numbers of men and women, and we all did the same jobs. I was assigned responsibility for all the computer elements of the food stamp program.

Because both money and legal requirements were involved, our work had to be exact. Errors that slipped through could have serious real-world consequences. We were trained to assume that every new program we wrote might have errors and to use various techniques to identify and fix them. We built test data sets that included every possible combination of unlikely events we could imagine, then ran them through our programs to make sure we got the desired results every time. We archived every version of our software. After each major change to the code, we would run the same data through the new and old versions of the program—sometimes for weeks—to make sure their results agreed exactly. When errors did pop up, we used the code archive to trace errors back to their origin and fix them at their source. We programmed the system to perform regular internal consistency checks as it ran. If any of those checks failed in the middle of the night, I had to get out of bed, drive down to the data center, and work as long as it took to find and fix the issue. In that programming job I developed a deep respect for data integrity and coding accuracy.

I also learned excellent programming skills for handling and reorganizing data—techniques I put to good use again and again throughout my scientific career.

The cavalier attitudes toward data and programming that I found later in research settings surprised me. Data, it seemed to me, were the currency of people's work, yet many people didn't understand how to manage data, didn't understand their tools, were poor programmers, and didn't test code adequately. In one high-profile case, a homemade data analysis program inherited from another lab was found to have an error that required the retraction of five papers, including three published in *Science* (56). Once, I noticed a study that had published results with the unusual finding of higher birth weights among female infants than male infants. Later, I happened to have access to the original data set and then saw that this finding had been an artifact caused by a coding problem. Years later, that result was noted as an unusual outlier in a comprehensive review of factors affecting birth weight (45). In fact, it was a coding error, but the reviewer could not have known that; indeed, the authors themselves probably never knew. I saw other examples in academia of problems with data management, coding, and record keeping. One group was unable to reproduce their own published results because they failed to record the code they had used. Another group knew almost nothing about the provenance of a data set that they were using; they didn't know how it had been created or even who had created it. In another study, there was some question about how the answers had been coded, but no one had kept a copy of the original data collection forms that would answer the question. Problems with software and data management in science persist but at least have begun to be recognized and addressed (54, 56, 79).

In criminal and civil law, proof of a documented, unbroken chain of custody is required for evidence to be considered admissible at a trial. I feel the same way about data—that it should be clear exactly where the data first came from and what changes may have been made along the way.

BACK TO CLASS

I was highly paid in my programming job and enjoyed the work, which was basically solving logic puzzles all day long. However, I didn't really like the black-box aspect and left after three years with the intention of studying nutrition, perhaps influenced by my work with the food stamp program. After community college courses in chemistry and biology, I reenrolled at UC to get a second bachelor's degree, this time in food and nutrition. While taking a summer class in organic chemistry, I found that I was pregnant. Our daughter, born the following year, turned out to have a very rare syndrome characterized by intractable seizures, brain malformations, and severe limitations. Her condition didn't get diagnosed until she was 5 years old. All that we knew at the time was that something was wrong, and we didn't know how wrong. I wanted to learn more about nutrition and applied to graduate programs in nutrition. Cornell University in Ithaca, New York, accepted my application and gave me funding. (I found out several years later that I had one of the highest Graduate Record Examination scores they had ever seen.)

By now we had a second child; we packed up the two children in winter and headed for Ithaca, where I had enrolled for a master's degree. That transition from Berkeley to Ithaca was challenging. I had never even seen snow on the ground. Now it was February in upstate New York. As I struggled across the snow-covered campus pushing a wheelbarrow containing a groggy sheep on which my physiology class had just performed a thyroidectomy, I wondered if I had done the right thing in leaving California.

I took classes in biochemistry, anthropology, psychology, economics, and international nutrition in a large and varied nutrition program, where numerous different disciplines were represented, with nutrition as the common theme. I had developed an interest in trace elements and chose as my advisor for my master's degree Wilson Pond, a professor in animal science, who supervised the project that became what would be my first contribution to the scientific literature. Working with pigs under Pond's expert guidance, I found that high dietary selenium reduced accumulation of cadmium in the liver, whereas dietary cadmium depressed tissue retention of added dietary selenium, probably by interfering with absorption. We published the results in the *Journal of Nutrition* (22).

Encouraged, I stayed on at Cornell for a doctorate. A newly arrived professor recommended that I take courses in mathematical statistics. My math background was so limited that I had to study a calculus text at home to keep up with the class. When, to my surprise, I received an A+ in a math-stat class, I belatedly realized I had an aptitude for math, and, unlike many people, I enjoyed statistics. I had also become interested in anthropometry, body composition, and the use of weight-height indices in epidemiologic studies. My supportive and helpful advisor for my doctorate was Jere Haas, a biological anthropologist. My dissertation combined my interests in population studies and epidemiology with my interest in anthropometry, misclassification, and measurement error. For this largely methodological study, I used data from the National Health Examination Survey (NHES), the predecessor of the National Health and Nutrition Examination Survey (NHANES). I couldn't have predicted it at the time, but I was to spend much of my subsequent research career involved with the NHANES.

My husband had followed me to Ithaca and then also enrolled in graduate school at Cornell, where we both finished our degrees in 1982. Now we were faced with a problem—finding jobs. Few entry-level academic positions were available. Many colleges and universities had recently ramped up hiring and filled up assistant professor slots. The University of Pittsburgh turned down my application for a postdoctoral fellowship in cardiovascular epidemiology. But then I got a surprise call offering me a postdoc there in alcohol epidemiology. It wasn't what I had had in mind, but it was a postdoc, so we moved to Pittsburgh. My spouse followed me to Pitt, where he did a postdoc in musicology.

I took additional statistics and epidemiology classes at Pitt and learned a lot about alcohol epidemiology, which had some interesting overlap with nutrition-related issues. However, it turned out that there was little opportunity to do research, and the program director expected me to spend part of my time working as the programmer for one of her projects. I declined as this hadn't been part of our agreement. I tried to make the best of the situation, earning a second master's degree (this one in public health), writing a thesis on issues with measuring alcohol consumption in epidemiological research, coauthoring a book chapter, and publishing a paper (18) based on my dissertation. As well, we had a third child.

UNDEREMPLOYED IN MICHIGAN

After these postdocs, my spouse got an appointment at the UM Society of Fellows, involving a three-year appointment as assistant professor in the music department and a concurrent three-year appointment as a postdoctoral scholar. Now I also needed to find something at UM. On my first visit to the campus to look for some opportunity, I was visibly pregnant, and the head of the nutrition program told me that this ruled me out for an appointment because they needed someone who could "hit the ground running." Finally, a statistician friend from Cornell told me that the National Institute of Child Health and Human Development (NICHD) had put out a request for proposals for studies of growth and nutrition that were designed for statisticians to be principal investigators. She suggested that I contact the UM biostatistics department to see if they were planning to bid on this proposal. They were writing a proposal and included me as a nutritionist on their grant application. It gradually became clear, however, that they saw me as a research assistant, not as an expert colleague. When NICHD posed several questions about their proposal, the statistician group wrote the answers to the nutrition questions themselves without consulting me, even though they had no expertise in nutrition. When NICHD asked the applicants to reduce the cost of their proposal, they cut me out of their application. Several years later, I discovered that one of the main reasons UM had failed to get the contract was because they no longer had a nutritionist involved. This seemed like poetic justice to me.

I got a job nonetheless in the biostatistics department—a half-time job to carry out analyses on a different grant with Richard (Dick) Landis, a leading survey statistician, and William (Bill) Harlan, an eminent cardiovascular epidemiologist. Landis first offered me a half-time research appointment. But, claiming expediency, he processed my job as a staff appointment, saying he would change it later. I learned a great deal from Harlan and Landis, two seasoned investigators who held their work to high standards. The purpose of the grant was to analyze secular trends in body mass index (BMI) categories for adolescents and young adults, using data from the four NHES/NHANES studies conducted between 1960 and 1980.

I had been hired to do programming and data analysis, but because I was also knowledgeable about the subject matter, I performed many of the functions of a coinvestigator on the project. I initiated analyses, wrote several of the articles that we produced (25, 26, 41), and even proposed the strategy that we used for modeling secular trends from repeated cross-sectional surveys. When the grant was renewed, I asked if I could become a coinvestigator but was turned down. At a conference several years later, Landis and two of his male grad students presented additional work they had done on the topic of modeling secular trends from repeated surveys, starting from the method I had proposed and displaying tables that I had generated. I was disappointed to see that I hadn't been invited to collaborate on the project and hadn't been credited for that work I had done.

I needed more income and wrote a grant proposal for the National Institute on Alcohol Abuse and Alcoholism (NIAAA). Unfortunately, I was not eligible to submit a proposal, because I was still classified as staff, despite several requests to get my status changed. To remedy this, I pointed out to the department chair that I had been told when I was hired that I would be switched to a research appointment. He balked but finally agreed to the switch, with the added caution that I had no future in the department. I was then able to submit my proposal and get my grant funded. However, I was clearly not considered in any way a colleague.

My time at Michigan was a fruitful period for me, nonetheless. I got a job participating in the analysis phase of a contract that Frances Larkin, a nutrition professor, had from the US Department of Agriculture (USDA) to see if a quantitative food frequency questionnaire (FFQ) could be used to evaluate nutrient intake in place of the three-day food records that the USDA was using. Sixteen days of dietary recall and records collected over 1 year were compared with estimates from a carefully designed 116-item FFQ administered in person (48). I devised a method of reexpressing food records into those components and then varying them systematically to see what component was causing disagreements between the FFQ and the food records (29, 30). Our analysis showed that frequency was the main cause of the discrepancies. We also found that disproportionate reporting of macronutrients limited the ability of FFQs to rank individuals similarly to a reference method, even if estimates were adjusted for total calorie intake. Several years later, I was invited to prepare a paper for a small International Life Sciences Institute workshop on the role of epidemiology in determining when evidence is sufficient to support nutrition recommendations. Drawing in part on the research I had done with Larkin, my paper (16) suggested that FFQ data were valuable for generating hypotheses but not adequate to provide evidence for or against specific recommendations. The other participants at the workshop, almost all of whom used FFQs regularly in their work, criticized my suggestion vehemently, but no one identified any errors in the paper except to say that it was "too negative."

My husband's three-year fellowship at UM was over, but there was a slight possibility that he might stay on there. Since I had no future in the biostatistics department (as the department chair had bluntly informed me), I tried to create a possible job for myself elsewhere at UM in case we ended up staying there. I now had a NIAAA grant paying half my salary and asked the chair of the community health services department, which included nutrition, to give a home to me and my grant. The chair asked for a copy of my curriculum vitae (CV). He then came to my office a few weeks later to tell me that he had circulated my CV to the faculty, who had been enthusiastic about taking me on. I could be an assistant research scientist in the department, he said, laying out what seemed to him an excellent offer. If all went well, I might be appointed as an assistant professor after four or five years. I thanked him politely and said I would let him know as our situation evolved. The prospect of perhaps finally becoming an assistant professor, 10 years after my doctorate, was depressing. But having the department chair think this was a great offer was so upsetting that I literally went home and cried. I felt like all my training in nutrition, statistics, and epidemiology had been useless. I had a PhD from an Ivy League university, had completed a two-year postdoc, and had worked for almost three years in the biostatistics department. Would the chair have thought this was an attractive offer for a man with similar credentials?

I had already seen what life was like at UM for an underclass of women with research appointments in the School of Public Health who supported themselves with grants. Departments might provide them an office in between grants to work on the next one, but no salary. In some cases, they retired during those hiatuses just to keep themselves funded and then unretired when a grant came through. I didn't want to join this marginalized group of women and was relieved not to stay at UM.

My husband was willing to follow me to another location, and I attempted to find a job. I applied for an entry-level faculty position at UC Davis. The three finalists for the position were one man, who was already an assistant professor elsewhere, and two women, including myself. Several months after my interview, I learned that they had offered the job to the man, who had declined. The department had decided that neither of the two women were qualified, so the position didn't get filled. Once again, I had been deemed unqualified for a job. I applied for an opening at the University of Minnesota. The department chair called to offer me the position, and they paid to fly my spouse out to talk with the music department about a position. While he was still in Minneapolis, they withdrew my job offer with no explanation. That was unfortunate because, on the basis of Minnesota's verbal offer, I had turned down an offer to join the Centers for Disease Control and Prevention (CDC)'s Epidemic Intelligence Service, which had a fixed cutoff date for acceptance. I ended up with no offers, but my husband got an offer from the Peabody Institute in Baltimore, Maryland, part of Johns Hopkins University. We decided to move to Baltimore, but I still needed to find work.

I BECOME A FEDERAL EMPLOYEE

I tried to sell my services to several universities in the Baltimore–Washington area, since I could pay half my salary with my own grant, but nobody was interested. While I was still at Michigan, I had been approached by Suzanne Haynes, who had come there to recruit applicants for her branch at the National Center for Health Statistics (NCHS), in the division that ran the NHANES. The ad published in the *American Journal of Epidemiology* (*AJE*) said, "Epidemiologists are being sought by the Division of Health Examination Statistics Program of NCHS to analyze data collected in the Hispanic Health and Nutrition Examination Surveys and previous Examination Surveys and to plan the Third National Health and Nutrition Examination Survey." I had never seriously considered working for the federal government but nevertheless applied for this position. Then, as now, government hiring proceeded at a glacial pace, and it didn't look like anything would come of this, but as we were moving to the Washington, DC, area, the NCHS position finally came through. The NCHS offered the advantages of a secure job, access to nutrition and anthropometric data, and an opportunity to do research and writing, which was what I really wanted to be doing. Plus, I was running out of options. By the time I was finally hired, months later, there had been several changes. The NCHS had been absorbed by the CDC, and Suzanne Haynes had left the agency.

My job at the NCHS turned into a career of almost 30 years, with many ups and downs. The federal government, at least at the NCHS and CDC, was a very different world from the academic environment that I had been used to. I was part of a hierarchical organization, and my activities were supposed to support the organization's policies and interests. I was surprised to learn that articles had to be cleared and approved by the agency before they could be submitted to a journal. I was assigned to a variety of projects that, although occasionally interesting, were outside my areas of expertise. For example, one of my assignments was to manage a large project digitizing 17,000 X-ray films of cervical and lumbar spines from NHANES II during the years 1976–80. Another time-consuming activity was giving descriptive presentations about different facets of the NHANES to various university and governmental audiences. I was sent to Michigan to give a talk on the rheumatology component of the NHANES and to New York to speak to a class at Columbia on the role of health examination surveys. I gave presentations to the NHANES Users Group and to the National Institutes of Health (NIH) Arthritis and Musculoskeletal Diseases Interagency Coordinating Committee. Despite being hampered by these and other multifarious activities, I continued to write articles and do analyses of NHANES data. In most cases I functioned as my own research assistant, doing literature searches, photocopying articles, doing all the programming and analyses, preparing my own tables and graphs, and doing whatever else needed to be done.

The NCHS was not the most ideal environment for research, but it had some advantages. As a statistical agency, it was neutral and not influenced by agendas. The emphasis was on statistical accuracy and data quality. Data-quality issues were particularly important in the NHANES, which was an examination survey with an in-person standardized examination and many laboratory measurements. Keeping measurements and procedures identical as much as possible across three different mobile examination centers and over six years or more was challenging. Laboratory quality control had to be consistently maintained for numerous analytes across a number of different contract labs. This large and complex survey required careful planning and administration and extensive oversight to run smoothly and accurately.

Data and analyses produced by the NCHS had important roles to play for many other government agencies. I had a front-row seat, and indeed considerable direct involvement, for many of the intricate struggles over matters such as FFQs, growth charts, and the risks of obesity—indeed even over the definition of obesity.

The NCHS started requiring all staff to attend diversity training in the 1990s. Unfortunately, as research has since shown, training alone has been largely ineffective at addressing institutional gender inequities (68). I saw that firsthand as I walked out of the diversity training session I attended. A male colleague turned to me and said, "So, now we have to take race and gender into account." Apparently, it had never occurred to this coworker that the preponderance of white males in the professions—as faculty members, as managers—might be precisely because their race and gender were being taken into account.

I experienced an example having to do with an editorial appointment. *AJE* was looking for new applicants for their long roster of associate editors. My supervisor suggested that both I and a male coworker with similar qualifications apply. A year later, I noticed that although I apparently hadn't been accepted, the male candidate's name now appeared in the published list. This caused me to realize for the first time that the list was approximately 95% males. This discovery led to a project in which I and four female colleagues from local universities gathered data from four influential epidemiology journals on their editors, reviewers, and authors. We found that women were significantly underrepresented among editors, not just overall but even relative to their contributions as reviewers and authors. We wrote up our findings in an article published in the *Journal of the American Medical Association* (*JAMA*) in 1998 (14a). Soon after our article was published—but almost eight years after I had originally applied—*AJE* invited me to be an associate editor. Despite my repeated requests to the editor, the journal almost never asked me to review articles in my areas of expertise. I had to wonder: Was I just a token female on the *AJE* roster?

I seemed to make some career progress at the NCHS. By 1991, I was promoted to chief of the medical statistics branch and served in this position for five years. Another woman was promoted to chief of the nutrition statistics branch. However, we were both in effect demoted when the male deputy division director created a new branch, made himself the chief of it, and then subsumed our branches into his. This unprecedented move allowed him several years later to bypass the one-year cooling-off period that government regulations would otherwise require before he took a new job that he had been offered by a contract organization. I kept my salary, but I lost my title and authority and no longer had anyone reporting to me. I was surprised that the CDC would allow two female branch chiefs to be so summarily removed through no fault of their own. I had seen male branch chiefs at the NCHS who had lost their jobs for some reason getting kicked upstairs into the division director's office. However, both I and the other woman, who had performed well and had not lost our positions for cause, were knocked down in the organizational chart to the same level at which I had started years earlier, now reporting to a branch chief, which was humiliating. We asked repeatedly to be at least assigned to the division director's office and were repeatedly rebuffed on the grounds that they didn't want too many people in that office, while at the same time several men were added to the office.

Primarily because of our scientific accomplishments, both she and I eventually became the most highly graded people in our division, attaining the title of distinguished consultants and a grade higher even than that of our division director. Yet within our division, we remained at the bottom of the organizational chart, at the lowest possible level of the hierarchy, unsupported and without any staff or organizational authority. This new and seemingly exalted title didn't come with any resources or assistance, so I still usually had to do much of the work myself, including not just writing, programming, and data analysis but also secretarial and administrative tasks, a pattern that continued across the rest of my career. I told my division director once that I should get paid less because 30% of my time was devoted to secretarial tasks.

CHARTING CHILDHOOD GROWTH

At the NCHS, I did a great deal of work related to childhood growth and growth charts (9, 10, 17, 31, 32, 34, 35, 47). The effort involved in creating new growth charts was challenging, but in addition, this proved to be a surprisingly contentious area. The NCHS itself had published the most widely used US growth charts in 1977 (40) and was executing a plan, developed before I arrived there, to revise the charts using data from NHANES III. Bob Kuczmarski, the NCHS staff member who was primarily responsible, was in close contact with the American Academy of Pediatrics and other researchers to discuss issues and keep them informed of the plans. The planning for NHANES III and the revision of the growth charts had been in the works for a long time, beginning in 1985.

The World Health Organization (WHO) had launched a different effort at a 1993 workshop (14). The WHO and NCHS intended to produce different kinds of charts. The new charts from the NCHS, to be called the CDC charts, would be descriptive of the US population; the

WHO's goal, in contrast, was to set prescriptive standards to be provided, along with dietary and breastfeeding recommendations, to member nations. Those standards would indicate how children should grow if they were full-term healthy singletons born to nonsmoking mothers, without socioeconomic constraints, not living at high altitude, and fed in line with WHO guidelines. This ambitious goal differed greatly from what the CDC was doing.

The two activities proceeded independently but began to collide. Advocates of the WHO approach went so far as to try to get the Secretary of Health and Human Services to forbid the CDC from publishing new charts, fearing that such charts would show breastfed babies as underweight. The CDC charts were released in 2000 (47) and the WHO charts a few years later (78a). The CDC held a workshop with participants from both sides and arrived at a compromise, probably unsatisfactory to all, that the CDC itself would recommend the use of the WHO charts up to age 2 and the CDC's own charts thereafter.

Publication of the new charts also entered into the continuing debate over the issue of overweight and obesity among children and the use of BMI to establish such categories. Although the WHO approach was supposed to generate data to show normal growth, WHO investigators arbitrarily excluded observations of heavier children that they considered to represent unhealthy weights for height. The CDC's analysis of normal growth, in contrast, included measurements on all children assessed. When I reanalyzed the CDC chart data using the same exclusions that the WHO had applied, the results turned out to be very similar to the WHO's standards (19).

In the United States, the whole topic of defining obesity and overweight for children was also contentious. Challenges were raised by a new expert committee (3) proposing new terminology for children. My colleague Cynthia Ogden and I argued unsuccessfully against new changes in nomenclature, in our view unjustified by any new data, that effectively labeled (and stigmatized) more children as overweight or obese (60, 66). I had concerns about whether BMI was even the best metric to use for children (15).

THE "OBESITY EPIDEMIC" BEGINS

Although the NHANES data gathered by the NCHS had often been used to assess overweight in the US population, the agency itself had yet to publish a journal article on that topic. To me this seemed like low-hanging fruit, something straightforward to write about that would be of general interest. Along with NCHS collaborators, I began working on an analysis that looked at trends from 1960 through NHANES III in 1988–91. At that point we were not using a category of obesity on the basis of BMI but only the category of overweight.

Using then-current definitions of overweight (a BMI of 27.8 or greater for men and 27.3 or greater for women, higher than those used today), we found that the prevalence had remained roughly the same from 1960 to 1980. But we were astonished to find an increase of almost 8 percentage points between 1976–80 and 1988–91. Previously, approximately a quarter of US adults were in the overweight category—had it really jumped to a third in just a decade? We treated the result with great caution and started checking every aspect of our data and analysis. A search for differences among the scales or the stadiometers at the three mobile examination centers turned up nothing. A review of the data processing found no anomalies. We scoured the literature for other articles that might have reported the same findings and casually asked colleagues at scientific meetings whether they thought the prevalence of overweight was changing. No one said they thought there was an increase.

We published the paper in 1994, with some trepidation (46). In the media and the scientific literature, this first clear indication that BMI values were beginning to rise attracted a lot of attention. The paper received almost 4,000 citations. Several people then told me that it had been obvious to them that an increase was happening, and similar results began to be reported in other countries.

Around the same time, new weight-loss drugs were coming onto the market. The companies that manufactured these products welcomed the idea of obesity as a burgeoning public health problem. In 1996, the US Food and Drug Administration (FDA) approved Redux (dexfenfluramine), the first weight-loss drug to make it to market since 1973 (12, 62). Just a year later, safety issues forced Redux to be withdrawn, but its appearance had demonstrated the huge demand for such products. The pharmaceutical industry began directing substantial money toward obesity research. The 1999 meeting of the North American Association for the Study of Obesity, previously a low-key affair, included an open bar, live music, ice sculptures, and a lavish banquet—all paid for by drug-company sponsorships.

New definitions of overweight and obesity were promulgated in the 1990s that made the estimates yet higher. The International Obesity Task Force (IOTF), funded primarily by drug companies (61, 67), was formed in 1996 for the purpose of drafting a report for a WHO consultation in 1997 (78). That report used the same BMI categories as had a 1995 WHO report (77) but renamed the categories. A BMI of 25 or above would be called overweight, and a BMI of 30 or above would now be obesity (78). This new terminology was adopted by the National Heart, Lung, and Blood Institute in 1998, thus in effect becoming US government policy. At the NCHS, we had no choice but to comply with this policy and with this terminology. I thought at the time that this was a good idea, because as a national statistical agency, it was preferable to use categories consistent with those in use at the NIH as well as in other countries and at the WHO.

The following year, a writer in the *New York Times* (73) observed that the new definitions had created "a booming new market for diet pills for the obese, practically served to the companies on a silver platter by the Government." I was shocked to see the silver platter statement and began to wonder whether these new categories might have been a mistake. The extent of the financing from drug companies was not clear to me at the time, but as IOTF chair Phil James described in a 2013 interview (67), "The people who funded the IOTF were drugs companies." And how much was he paid? "They used to give me cheques for about 200,000 [British pounds] a time. And I think I had a million or more." And did they ever ask him to push any specific agenda? "Not at all."

The new definitions labeled most of the US population as above normal weight. In 1998, the American Heart Association (AHA) reclassified obesity as a major risk factor for heart disease. The AHA (11) and other public health interest groups warned of dire health consequences. "Obesity is the terror within," said Surgeon General Richard Carmona (71, p. 145). "Unless we do something about it, the magnitude of the dilemma will dwarf 9/11 or any other terrorist attempt." Other warnings of dire health consequences reverberated from medical and government authorities through popular culture.

I began to get the feeling that what was going on was excessive and exaggerated somehow. Had it become, as a 2005 article in *Scientific American* suggested (38), an overblown epidemic? On the one hand, above-average weight had been associated with increased mortality and increased risk of factors such as elevated blood pressure and dyslipidemia (43, 72). On the other hand, almost half of the US population in 1960 had been above what we were now calling normal weight. As that proportion grew in the following decades, overall health improved, life expectancies increased, and cardiovascular mortality dropped steadily. By 2000, age-adjusted mortality rates from coronary heart disease and stroke had declined by approximately two-thirds from 1960s levels (53). I found it disturbing to see discussions of obesity laced with inherently racist, sexist, and classist comments, with poor Black women sometimes being singled out.

My research explored different aspects of the changes in definitions and the effect of those changes and other factors on trends in the prevalence of overweight and obesity. I published numerous papers on these questions, including one in the *New England Journal of Medicine* on the possible effect of smoking cessation (33). I continued to publish methodological papers, including an article showing a counterintuitive effect of categorizing a baseline exposure measurement (27).

I also continued to publish papers on the ongoing changes in BMI. The NHANES became a continuous survey in 1999, with new data released every two years. Along with Cynthia Ogden, also at the NCHS, I wrote a series of papers about trends in the prevalence of various BMI categories. Our first such paper (21) received almost 4,600 citations, and the next paper (42) had more than 3,000. We continued to publish articles in *JAMA* for each additional year cycle (20, 65). We developed a good relationship with *JAMA*, partly because these articles were so highly cited and also because the journal trusted our work to be accurate. We stuck closely to the facts without editorializing. As representatives of a statistical agency, our job was to present the data, not to engage in public health messaging.

THE OBESITY WARS

One concern was how many deaths per year could be caused by obesity, a murky question at best. One estimate of 300,000 deaths per year in the United States had gained significant traction. It was cited as part of the FDA's approval process for Redux, which not long thereafter turned out to have serious safety issues and was withdrawn from the market (62). This number had been derived from a 1993 article by McGinnis & Foege (51), who subsequently wrote a letter to the *New England Journal of Medicine* in 1998 (52) clarifying that their estimate was not an estimate of deaths caused by obesity.

In 1998, I joined forces with David Williamson, a fellow nutritionist at the CDC whom I knew from graduate school at Cornell, to try to use NHANES data to estimate obesity-associated deaths. The main problem was that we couldn't figure out how best to do this. An innovative approach had been proposed by Allison et al. in 1999 (2), but we wanted to improve on the method and to use only NHANES data. We enlisted Barry Graubard from the National Cancer Institute (NCI), an expert survey statistician, to discuss methodological approaches. It was a lengthy project that required grappling with a lot of statistical and data issues, many of which we laid out in a 2004 article and commentary (23, 36). The key contribution came from Mitchell Gail, a leading statistician and medical doctor at NCI who joined our collaboration and suggested we use a modification of a method he had developed to assess individuals' risk of developing breast cancer (37). We set to work.

We were surprised to see an article on the same topic in the March 2004 issue of *JAMA* with Ali Mokdad, then chief of the CDC's behavioral surveillance branch, as first author and CDC Director Julie Gerberding as senior author (57). The authors estimated that overweight and obesity were associated with 500,000 deaths in 2000 and projected that obesity could soon overtake smoking as a cause of death. The Mokdad et al. paper generated huge media coverage. Both the estimate and the comparison with tobacco deaths perturbed tobacco-control researchers and caused confusion and chaos both within and outside the CDC (13, 50). A congressman requested a General Accounting Office (GAO) investigation, which prompted the CDC to replicate the analyses. The CDC reported to the GAO that they had been able to replicate the values for overweight and obesity perfectly.

Though we were still working on our paper, it was already clear from the analyses we had completed that our estimates were going to be much lower. Comparisons of our numbers with the estimates of Mokdad et al. were inevitable, so we decided we should reproduce the Mokdad et al. estimates and then sequentially adjust for differences. All the information needed to reproduce the Mokdad et al. estimates had already been published.

To our astonishment, we could not replicate the Mokdad et al. estimates. When Williamson asked some polite questions about how the calculations had been done, he was told he was harassing the authors and was forbidden to contact them further. We finally obtained a copy of the calculations Mokdad et al. had performed. We knew from our own work that accurate estimates required a sophisticated analysis to account for many potentially confounding factors. Williamson and I had even talked with Mokdad several times and explained the methodological difficulties involved in making estimates of this kind. But the method in the paper by Mokdad et al. was extremely simple: just 10 lines in an Excel spreadsheet.

Even more disturbing, a cursory inspection of the spreadsheet revealed that the authors had made major errors, including using the vital statistics numbers of deaths for the wrong year in five out of six cases in the spreadsheet and using the wrong prevalence of BMI <25, causing the sum of BMI categories to be greater than 100%. This was a remarkable example of a careless approach to data and coding in a highly visible publication. The CDC passed the errors off as a computer glitch, rather than the result of mistakes in data handling and computations (57, 58).

Meanwhile, we had submitted our paper to *JAMA*. Instead of the 500,000 deaths in 2000 associated with obesity reported by Mokdad et al., we found approximately 26,000 net excess deaths: 112,000 excess deaths associated with obesity minus 86,000 fewer deaths than expected for the overweight category. *JAMA* accepted our paper and published it promptly in April 2005 (24). The lower mortality in the overweight category was hardly a new observation; the 1998 clinical guidelines on overweight and obesity put out by the NIH (63) had noted that "Recently, a 20-year prospective study of a nationally representative sample of U.S. adults age 55 to 74 years suggested that lowest mortality occurs in the BMI range of 25 to 30."

Shortly before our article was published, it finally began to dawn on CDC upper management that the discrepancies between our findings and those of Mokdad et al. might be a problem. The source of the differences was not mysterious: We had used better methods and better data. But I began to receive stern and unfriendly phone calls from a high-level staffer in the CDC director's office, far above me in the hierarchy. He delivered marching orders from the top: Don't say that our results were better or that we had used better data and better methods; instead say that we had gotten different results because we had used different data and different methods.

Concerned about possible repercussions, I complied. A CDC press official stayed on the line for every call I had with media, listening to the questions and my responses. The media interest was so intense I had to use a separate phone line to field interviews. But when reporters and others asked me why our results were different, I had to leave the incongruity mysteriously unexplained. Even to this day, an impression persists in some quarters that both the Mokdad et al. estimates and our estimates were reasonably valid or that we had simply reanalyzed the same data that Mokdad et al. had used. Though Gerberding gave a press conference several months later in response to internal pressure, her statements seemed designed to throw shade on our paper. "What we don't know yet is the best methods for estimating the impact of obesity on death," she said. "And we are committed to learning more about that process, but you can expect that more studies will be coming forward and that more information will be made available as this research is conducted" (7a). The next day, a CDC press official called to apologize to me and said that those remarks were not among the talking points they had prepared for the director. Despite all the furor within the CDC, the agency eventually accepted our estimates as the more correct numbers.

DATA AND ITS INTERPRETATION

The backlash and controversy over our 2005 paper were astonishing. When in 1994 I published a paper showing that overweight was increasing, these findings had been quickly accepted, and indeed welcomed as a call to action, by the public health world. I was invited to present my work widely in the United States and in international venues. In the 2005 paper, we had found significantly fewer deaths in the overweight category than expected relative to normal weight. These findings had been treated as bizarre and as being completely different from some imagined consensus that overweight led to increased mortality risk. I have described elsewhere how I was subjected to years of bullying and ad feminam attacks by faculty and alumni of the Harvard School of Public Health (17a) after the publication of our article. They and others seemed determined to sow doubts about our results. Our paper was straightforward and defensible, used only publicly available data, and corrected the errors in several previous papers on the topic. The NCHS was very strict about forbidding articles to make any recommendations for action or any policy statements, and our article contained no such statements. Nonetheless, some were concerned because our findings might possibly be hijacked by special interest groups to influence policy makers, although other scientists were uncomfortable with the idea of hiding or dismissing data for the sake of a simpler message (44).

I went back to the literature and realized that our results were consistent with much of the previous research on the issue. I decided to address the problem in a systematic way through a meta-analysis carried out with researchers from the NCHS, the NCI, and the University of Ottawa. We identified all studies of BMI and mortality in the general adult population that had used the same standard BMI categories that we had used (28). This involved reviewing almost 7,000 articles. We found 97 such studies, with almost 3 million people, and observed that the results were consistent with those in our 2005 paper. But in many of those 97 studies, those findings—perhaps because they seemed contrary to the prevailing narrative—had been buried in the tables or in the appendix, not highlighted in the discussion and not mentioned in the abstract.

Over the years I had built up a good reputation in the obesity research community as a careful researcher who was an accurate reporter and didn't go beyond the data. Now I realized that part of this was because people liked the results that we got, especially the increasing prevalence of obesity. These results were greeted uncritically. Now we were following the same careful approach in our 2005 and 2013 papers, yet our results were apparently far less desirable. My training and background inclined me to follow analyses wherever they might lead, instead of trying to match an expected outcome or a public health message. This approach was not always welcome.

LESS THAN HOPED, BUT MORE THAN EXPECTED

Several things along the way have shaped my outlook and formed aspects of my professional career. I had learned from my programming job the importance of doing things right and keeping track of everything, of testing and double-checking. I had also learned to be an excellent programmer myself. I landed almost by accident at the NCHS, which was not a public health or research agency per se, but rather a federal statistical agency whose job it was to make factual statements. I was unable to get a university job and indeed also failed to be hired for several jobs in other federal agencies. I made little career progress. Looking back, I realize it is likely that some of the challenges I faced might have been related to aspects of unconscious bias. Although I was aware of research showing that, for example, a male name instead of a female name on a resume for an applicant would lead to a higher salary offer and a better evaluation of competence (59), I really didn't think of this as applying to me personally. But perhaps if the name on my CV had been Kevin or Kenneth instead of Katherine I would have been judged more favorably and offered more opportunities. Orchestras began to have more female players when blind auditions began to be conducted that hid the performer's gender (39). Though overt sexism in science and workplaces has diminished notably over the course of my career, progress toward eliminating unconscious bias has been slower. The Matthew and Matilda effects are still powerful forces that distort incentives and rewards in science (6, 7, 68, 70, 76). I have no real way of knowing how much this affected my career, but I did often feel undervalued and underrecognized. I could never get a faculty job and was often treated like an underling.

I would have liked to have been able to do more research, as I always had many more original ideas than I could execute, but the lack of resources was a limiting factor. I could have accomplished more if I had had more resources available, but all told, I did better and had more of an impact than I would ever have thought. It was lucky that I had few aspirations because I got few rewards. My interest was and remains in doing good work and answering questions or at least shedding light on questions that others have not always considered carefully. Other than that, I ended up with a not atypical female career, striving to do good work for its own sake and receiving relatively little recognition from it.

Sometimes recognition comes in unexpected forms, even after many years. At the NCI, I was once introduced in the elevator to a leading researcher there. He heard my name and asked if I had written "that article." I didn't know what article he meant. He turned out to be referring to an article on the effects of exposure misclassification on estimates of relative risk (18). This was the second article I ever published, literally decades before. The article had gotten little recognition, and this was the first time anyone had ever mentioned it to me. He said that the article had been extremely valuable for his lab group because it made them realize that they needed to improve the sensitivity of their analyses, not the specificity. I have heard from many other people who have thanked me for a variety of aspects of my work and told me that it was valuable for their teaching and research, so I am proud to say that my work did make some real contributions in both methodological and subject areas.

DISCLOSURE STATEMENT

The author is not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

For editorial assistance, I would like to thank Wayt Gibbs and John Spitzer, my ever-supportive husband. I also thank Walter Willett for his role in drawing so much attention to my work.

LITERATURE CITED

1. Aguinis H, Ji YH, Joo H. 2018. Gender productivity gap among star performers in STEM and other scientific fields. *J. Appl. Psychol.* 103:1283–306
2. Allison DB, Fontaine KR, Manson JE, Stevens J, VanItallie TB. 1999. Annual deaths attributable to obesity in the United States. *JAMA* 282:1530–38
3. Barlow SE. 2007. Expert committee recommendations regarding the prevention, assessment, and treatment of child and adolescent overweight and obesity: summary report. *Pediatrics* 120(Suppl. 4):S164–92
4. Barres BA. 2006. Does gender matter? *Nature* 442:133–36
5. Barres BA. 2018. *The Autobiography of a Transgender Scientist*. Cambridge, MA: MIT Press
6. Begeny CT, Ryan MK, Moss-Racusin CA, Ravetz G. 2020. In some professions, women have become well represented, yet gender bias persists—perpetuated by those who think it is not happening. *Sci. Adv.* 6:eaba7814
7. Casad BJ, Franks JE, Garasky CE, Kittleman MM, Roesler AC, et al. 2021. Gender inequality in academia: problems and solutions for women faculty in STEM. *J. Neurosci. Res.* 99:13–23

7a. CDC (Cent. Dis. Control Prev.). 2005. *Overweight and obesity: clearing the confusion.* Telebriefing Transcript, June 2. **https://www.cdc.gov/media/transcripts/t050602.htm**

8. Chan HF, Torgler B. 2020. Gender differences in performance of top cited scientists by field and country. *Scientometrics* 125:2421–47

9. Cole TJ, Bellizzi MC, Flegal KM, Dietz WH. 2000. Establishing a standard definition for child overweight and obesity worldwide: international survey. *BMJ* 320:1240–43
10. Cole TJ, Flegal KM, Nicholls D, Jackson AA. 2007. Body mass index cut offs to define thinness in children and adolescents: international survey. *BMJ* 335:194
11. Collins G. 2009. *When Everything Changed*. New York: Little, Brown & Company
12. Colman E. 2005. Anorectics on trial: a half century of federal regulation of prescription appetite suppressants. *Ann. Intern. Med.* 143:380–85
13. Couzin J. 2005. Public health. A heavyweight battle over CDC's obesity forecasts. *Science* 308:770–71
14. de Onis M, Garza C, Victora CG, Onyango AW, Frongillo EA, Martines J. 2004. The WHO Multicentre Growth Reference Study: planning, study design, and methodology. *Food Nutr. Bull.* 25:S15–26

14a. Dickersin K, Fredman L, Flegal KM, Scott JD, Crawley B. 1998. Is there a sex bias in choosing editors? Epidemiology journals as an example. *JAMA* 280:260–64

15. Flegal KM. 1999. Curve smoothing and transformations in the development of growth curves. *Am. J. Clin. Nutr.* 70:163S–65S
16. Flegal KM. 1999. Evaluating epidemiologic evidence of the effects of food and nutrient exposures. *Am. J. Clin. Nutr.* 69:1339S–44S
17. Flegal KM. 2000. The effects of age categorization on estimates of overweight prevalence for children. *Int. J. Obes. Relat. Metab. Disord.* 24:1636–41

17a. Flegal KM. 2021. The obesity wars and the education of a researcher: a personal account. *Prog. Cardiovasc. Dis*. 67:75–79

18. Flegal KM, Brownie C, Haas JD. 1986. The effects of exposure misclassification on estimates of relative risk. *Am. J. Epidemiol.* 123:736–51
19. Flegal KM, Carroll MD, Ogden CL. 2012. Effects of trimming weight-for-height data on growth-chart percentiles. *Am. J. Clin. Nutr.* 96:1051–55
20. Flegal KM, Carroll MD, Ogden CL, Curtin LR. 2010. Prevalence and trends in obesity among US adults, 1999–2008. *JAMA* 303:235–41
21. Flegal KM, Carroll MD, Ogden CL, Johnson CL. 2002. Prevalence and trends in obesity among US adults, 1999–2000. *JAMA* 288:1723–27
22. Flegal KM, Cary EE, Pond WG, Krook LP. 1980. Dietary selenium and cadmium interrelationships in growing swine. *J. Nutr.* 110:1255–61
23. Flegal KM, Graubard BI, Williamson DF. 2004. Methods of calculating deaths attributable to obesity. *Am. J. Epidemiol.* 160:331–38
24. Flegal KM, Graubard BI, Williamson DF, Gail MH. 2005. Excess deaths associated with underweight, overweight, and obesity. *JAMA* 293:1861–67
25. Flegal KM, Harlan WR, Landis JR. 1988. Secular trends in body mass index and skinfold thickness with socioeconomic factors in young adult men. *Am. J. Clin. Nutr.* 48:544–51
26. Flegal KM, Harlan WR, Landis JR. 1988. Secular trends in body mass index and skinfold thickness with socioeconomic factors in young adult women. *Am. J. Clin. Nutr.* 48:535–43
27. Flegal KM, Keyl PM, Nieto FJ. 1991. Differential misclassification arising from nondifferential errors in exposure measurement. *Am. J. Epidemiol.* 134:1233–44
28. Flegal KM, Kit BK, Orpana H, Graubard BI. 2013. Association of all-cause mortality with overweight and obesity using standard body mass index categories: a systematic review and meta-analysis. *JAMA* 309:71–82
29. Flegal KM, Larkin FA. 1990. Partitioning macronutrient intake estimates from a food frequency questionnaire. *Am. J. Epidemiol.* 131:1046–58
30. Flegal KM, Larkin FA, Metzner HL, Thompson FE, Guire KE. 1988. Counting calories: partitioning energy intake estimates from a food frequency questionnaire. *Am. J. Epidemiol.* 128:749–60
31. Flegal KM, Ogden CL, Wei R, Kuczmarski RL, Johnson CL. 2001. Prevalence of overweight in US children: comparison of US growth charts from the Centers for Disease Control and Prevention with other reference values for body mass index. *Am. J. Clin. Nutr.* 73:1086–93
32. Flegal KM, Tabak CJ, Ogden CL. 2006. Overweight in children: definitions and interpretation. *Health Educ. Res.* 21:755–60

33. Flegal KM, Troiano RP, Pamuk ER, Kuczmarski RJ, Campbell SM. 1995. The influence of smoking cessation on the prevalence of overweight in the United States. *N. Engl. J. Med.* 333:1165–70
34. Flegal KM, Wei R, Ogden C. 2002. Weight-for-stature compared with body mass index-for-age growth charts for the United States from the Centers for Disease Control and Prevention. *Am. J. Clin. Nutr.* 75:761–66
35. Flegal KM, Wei R, Ogden CL, Freedman DS, Johnson CL, Curtin LR. 2009. Characterizing extreme values of body mass index for age by using the 2000 Centers for Disease Control and Prevention growth charts. *Am. J. Clin. Nutr.* 90:1314–20
36. Flegal KM, Williamson DF, Pamuk ER, Rosenberg HM. 2004. Estimating deaths attributable to obesity in the United States. *Am. J. Public Health* 94:1486–89
37. Gail MH, Brinton LA, Byar DP, Corle DK, Green SB, et al. 1989. Projecting individualized probabilities of developing breast cancer for white females who are being examined annually. *J. Natl. Cancer Inst.* 81:1879–86
38. Gibbs WW. 2005. Obesity: an overblown epidemic? *Sci. Am.* 292:70–77
39. Golden C, Rouse C. 2000. Orchestrating impartiality: the impact of "blind" auditions on female musicians. *Am. Econ. Rev.* 90:715–41
40. Hamill PV, Drizd TA, Johnson CL, Reed RB, Roche AF. 1977. NCHS growth curves for children birth–18 years. United States. *Vital. Health Stat.* 11:i–74
41. Harlan WR, Landis JR, Flegal KM, Davis CS, Miller ME. 1988. Secular trends in body mass in the United States, 1960–1980. *Am. J. Epidemiol.* 128:1065–74
42. Hedley AA, Ogden CL, Johnson CL, Carroll MD, Curtin LR, Flegal KM. 2004. Prevalence of overweight and obesity among US children, adolescents, and adults, 1999–2002. *JAMA* 291:2847–50
43. Hubert HB, Feinleib M, McNamara PM, Castelli WP. 1983. Obesity as an independent risk factor for cardiovascular disease: a 26-year follow-up of participants in the Framingham Heart Study. *Circulation* 67:968–77
44. Hughes V. 2013. The big fat truth. *Nature* 497:428–30
45. Kramer MS. 1987. Determinants of low birth weight: methodological assessment and meta-analysis. *Bull. World Health Organ.* 65:663–737
46. Kuczmarski RJ, Flegal KM, Campbell SM, Johnson CL. 1994. Increasing prevalence of overweight among US adults. The National Health and Nutrition Examination Surveys, 1960 to 1991. *JAMA* 272:205–11
47. Kuczmarski RJ, Ogden CL, Guo SS, Grummer-Strawn LM, Flegal KM, et al. 2002. 2000 CDC growth charts for the United States: methods and development. *Vital. Health Stat.* 11:1–190
48. Larkin FA, Metzner HL, Thompson FE, Flegal KM, Guire KE. 1989. Comparison of estimated nutrient intakes by food frequency and dietary records in adults. *J. Am. Diet. Assoc.* 89:215–23
49. Lincoln AE, Pincus S, Koster JB, Leboy PS. 2012. The Matilda effect in science: awards and prizes in the US, 1990s and 2000s. *Soc. Stud. Sci.* 42:307–20
50. Marshall E. 2004. Epidemiology. Public enemy number one: tobacco or obesity? *Science* 304:804
51. McGinnis JM, Foege WH. 1993. Actual causes of death in the United States. *JAMA* 270:2207–12
52. McGinnis JM, Foege WH. 1998. The obesity problem. *N. Engl. J. Med.* 338:1157
53. Mensah GA, Wei GS, Sorlie PD, Fine LJ, Rosenberg Y, et al. 2017. Decline in cardiovascular mortality: possible causes and implications. *Circ. Res.* 120:366–80
54. Merali Z. 2010. Computational science: . . .Error. *Nature* 467:775–77
55. Merton RK. 1968. The Matthew effect in science. The reward and communication systems of science are considered. *Science* 159:56–63
56. Miller G. 2006. Scientific publishing. A scientist's nightmare: Software problem leads to five retractions. *Science* 314:1856–57
57. Mokdad AH, Marks JS, Stroup DF, Gerberding JL. 2004. Actual causes of death in the United States, 2000. *JAMA* 291:1238–45
58. Mokdad AH, Marks JS, Stroup DF, Gerberding JL. 2005. Correction: actual causes of death in the United States, 2000. *JAMA* 293:293–94
59. Moss-Racusin CA, Dovidio JF, Brescoll VL, Graham MJ, Handelsman J. 2012. Science faculty's subtle gender biases favor male students. *PNAS* 109:16474–79

60. Moynihan R. 2006. Expanding definitions of obesity may harm children. *BMJ* 332:1412
61. Moynihan R. 2006. Obesity task force linked to WHO takes “millions” from drug firms. *BMJ* 332:1412
62. Mundy A. 2001. *Dispensing with the Truth: The Victims, the Drug Companies, and the Dramatic Story Behind the Battle over Fen-Phen*. New York: St. Martin’s Press
63. NIH (Natl. Inst. Health). 1998. Clinical guidelines on the identification, evaluation, and treatment of overweight and obesity in adults—the evidence report. National Institutes of Health. *Obes. Res.* 6(Suppl. 2):51S–209S
64. Nordell J. 2021. *The End of Bias: A Beginning: The Science and Practice of Overcoming Unconscious Bias*. New York: Metropolitan Books
65. Ogden CL, Carroll MD, Curtin LR, McDowell MA, Tabak CJ, Flegal KM. 2006. Prevalence of overweight and obesity in the United States, 1999–2004. *JAMA* 295:1549–55
66. Ogden CL, Flegal KM. 2010. *Changes in terminology for childhood overweight and obesity*. Natl. Health Stat. Rep. 25, Natl. Cent. Health Stat., Hyattsville, MD
67. Peretti J. 2013. Fat profits: how the food industry cashed in on obesity. *The Guardian*, Aug. 7. **https://www.theguardian.com/lifeandstyle/2013/aug/07/fat-profits-food-industry-obesity**
68. Pritlove C, Juando-Prats C, Ala-Leppilampi K, Parsons JA. 2019. The good, the bad, and the ugly of implicit bias. *Lancet* 393:502–4
69. Rossiter MW. 1993. The Matthew Matilda effect in science. *Soc. Stud. Sci.* 23:325–41
70. Sá C, Cowley S, Martinez M, Kachynska N, Sabzalieva E. 2020. Gender gaps in research productivity and recognition among elite scientists in the U.S., Canada, and South Africa. *PLOS ONE* 15:e0240903
71. Saguy A. 2013. *What’s Wrong with Fat?* New York: Oxford University Press
72. Sorlie P, Gordon T, Kannel WB. 1980. Body build and mortality. The Framingham study. *JAMA* 243:1828–31
73. Stolberg SG. 1999. Ideas & trends: The fat get fatter; overweight was bad enough. *New York Times*, May 2. **https://www.nytimes.com/1999/05/02/weekinreview/ideas-trends-the-fat-get-fatter-overweight-was-bad-enough.html**
74. Thompson C. 2019. *Coders: The Making of a New Tribe and the Remaking of the World*. New York: Penguin Press
75. Thompson C. 2019. The secret history of women in coding. *New York Times*, Feb. 13. **https://www.nytimes.com/2019/02/13/magazine/women-coding-computer-programming.html**
76. Vásárhelyi O, Zakhlebin I, Milojević S, Horvát E. 2021. Gender inequities in the online dissemination of scholars’ work. *PNAS* 118:e2102945118
77. WHO (World Health Organ.). 1995. Physical status: the use and interpretation of anthropometry. Report of a WHO Expert Committee. *World Health Organ. Tech. Rep. Ser.* 854:1–452
78. WHO (World Health Organ.). 2000. Obesity: preventing and managing the global epidemic. Report of a WHO consultation. *World Health Organ. Tech. Rep. Ser.* 894:i–253

78a. WHO (World Health Organ.). 2006. *WHO child growth standards: length/height-for-age, weight-for-age, weight-for-length, weight-for-height and body mass index-for-age: methods and development.* Rep., World Health Organ., Geneva. **https://apps.who.int/iris/bitstream/handle/10665/43413/924154693X_eng.pdf**

79. Wilson G, Bryan J, Cranston K, Kitzes J, Nederbragt L, Teal TK. 2017. Good enough practices in scientific computing. *PLOS Comput. Biol.* 13:e1005510