Between Bench and Bedside: Building Clinical Consensus at the NIH, 1977–2013

TODD M. OLSZEWSKI

Health Policy and Management Department, Providence College, Providence, Rhode Island, 02918.

Email: tolszewski@providence.edu.

ABSTRACT

After World War II, the National Institutes of Health (NIH) emerged as a major patron of biomedical research. In the succeeding decades, NIH administrators sought to determine how best to disseminate the findings of the research it supported and manage their relationship with clinicians in the national community. This task of bridging research and practice fell to the Office of Medical Applications of Research (OMAR), which administered the NIH Consensus Development Program (CDP) between 1978 and 2012. This article argues that the CDP represented an unusual attempt to depoliticize biomedical research and medical practice at a particularly controversial time in American medicine. Throughout the program's existence, administrators sought ways to bring new knowledge to the medical community without creating the appearance of regulating clinical practice. For an agency with a mandate to promote the production of new biomedical knowledge, the question remained open as to how far this responsibility extended from the bench to the bedside. In striking this balance, the leadership sought to refine their understanding of the role and mission of the NIH. The history of the CDP has much to tell us about postwar biomedical research, health politics, and the institutional development of the NIH.

KEYWORDS: National Institutes of Health, biomedical research, consensus conferences, evidence-based medicine, translational medicine

With the rise of “big science” during World War II came “big medicine.” Formally founded by the Ransdell Act in 1930, the National Institutes of Health (NIH) emerged as the major patron of biomedical research in the United States after the Second World War.¹ Comprised of ten institutes by 1968, the NIH was responsible for sponsoring

1 Caroline Hannaway, ed., *Biomedicine in the Twentieth Century: Practices, Policies, and Politics* (Amsterdam: IOS Press, 2008); Stephen Strickland, *Politics, Science, and Dread Disease: A Short History of United States Medical Research Policy* (Cambridge: Harvard University Press, 1972).

basic and clinical research both at its Maryland campus and throughout the country. Under the leadership of director James Shannon, the NIH formed several new institutes and the created the Clinical Center as the agency's flagship research hospital, as NIH appropriations increased fourteen-fold between 1950 and 1968.² By 1974, the NIH was funding almost half of all biomedical research and development at American medical schools and universities.³ By the time that Donald Fredrickson became the director in 1975, the NIH had settled into relative organizational stability.⁴

Public sentiments about science and the government were a different matter. In the 1970s, even before Watergate, Americans had become distrustful of their government. For almost twenty years, the country had been mired in the Vietnam War; the president was under investigation for political crimes; and the War on Drugs had demonized poor and minority communities across the nation. The American medical profession, long held in high regard, also met with criticism and distrust. The revelations in 1972 that a government-sponsored study had prevented hundreds of African American men from obtaining treatment for syphilis undermined trust in the health care system.⁵ The scandals surrounding the Dalkon Shield intrauterine device in the early 1970s shook confidence in the promise of new medical technologies.⁶ Women united against the orthodox medical profession to demand more information about and control over their bodies.⁷ Physicians seemed ever more distant as hospital administrators and federal bureaucrats codified and depersonalized their work.⁸ Meanwhile, health care costs seemed to rise unabated.⁹

It was in this environment that NIH administrators, perhaps naively, sought to expand the reach of their institution to offer more guidance to practicing physicians and the broader public. In 1978, the NIH opened the Office of Medical Applications of Research (OMAR) to communicate knowledge generated by NIH-funded research to the medical community and the public.¹⁰ Until its dissolution in 2013, the OMAR

- 2 Shannon served as NIH Director from 1955-1968. In 1950, NIH appropriations were \$52,710,000; by 1968, NIH appropriations were \$1,076,000,000 (\$745,160,919 in 1950 dollars). For the CPI inflation calculator, see <https://data.bls.gov/cgi-bin/cpicalc.pl>. For NIH appropriations data, see National Institutes of Health, Office of Budget, "Appropriations History by Institute/Center (1938-present)," https://ofofbudget.od.nih.gov/approp_hist.html.
- 3 National Institutes of Health, "Basic Data Relating to the National Institutes of Health" (Bethesda, MD: Office of Program Planning & Evaluation and the Division of Research Grants, 1976): 6.
- 4 Even though the rate of growth of NIH funding outpaced the rate of growth of overall federal funding for research and development activities, NIH appropriations increased less than two-fold during the 1970s. See also, Office of Management and Budget, "Historical Tables," Table 9.8, <https://www.whitehouse.gov/omb/budget/Historicals>.
- 5 Susan Reverby, *Examining Tuskegee: The Infamous Syphilis Study and Its Legacy* (Chapel Hill: University of North Carolina Press, 2009).
- 6 Andrea Tone, *Devices and Desires: A History of Contraceptives in America* (New York: Hill and Wang, 2001).
- 7 Wendy Kline, *Bodies of Knowledge: Sexuality, Reproduction, and Women's Health in the Second Wave* (Chicago: University of Chicago Press, 2010).
- 8 David J. Rothman, *Strangers at the Bedside: A History of How Law and Bioethics Transformed Medical Decision Making* (New York: Basic Books, 1991).
- 9 William Baumol, *The Cost Disease: Why Computers Get Cheaper and Health Care Doesn't* (New Haven: Yale University Press, 2012).
- 10 "Organizational Structure of the National Institutes of Health, Background Paper for the NAS Committee on the Organizational Structure of the NIH" (McGeary and Smith, Washington, D.C., June 20, 2002), 76.

operated the Consensus Development Program (CDP), which held conferences on critical and often controversial topics, bringing together panels of experts to exchange opinions and create official “consensus statements.”

This article examines this history of the CDP as a critical moment in late twentieth-century medicine. I argue that the CDP represented an attempt by the federal government to depoliticize biomedical research and medical practice at a particularly controversial time in American medicine. NIH administrators hoped not only to depoliticize biomedicine but also to establish new models for consensus even as the politics of identity and waves of social activism polarized the nation. By attempting to insert the NIH into questions of medical practice, administrators were challenging the traditional role of the agency as a site and support for the production of biomedical knowledge. Could the government find a place near—but not quite at—the bedside?

While the NIH's postwar rise as a major patron of biomedical research has been well documented, historians have just begun to examine NIH administrators' efforts to interpret the NIH mission in relation to efforts to bridge scientific research and clinical practice.¹¹ This article uses the history of OMAR and the CDP to consider how NIH administrators sought to disseminate the abundant findings of the research it supported and manage their relationship with clinicians working in hospitals and clinics, often far from the laboratory.¹² During almost four decades, the CDP produced 159 consensus statements on topics as varied as acupuncture, gastric bypass surgery, total hip replacement, and the health benefits of pets (full listing available at (<https://consensus.nih.gov/previous.htm>) and in Appendix A) (Figure 1). Historians of medicine have examined specific CDP conferences related to debates about cholesterol

11 On NIH's intramural research program, see Victoria Harden, *Inventing the NIH: Federal Biomedical Research Policy, 1887-1937* (Baltimore: Johns Hopkins University Press, 1986); Buhm Soon Park, “The Development of the Intramural Research Program at the National Institutes of Health after World War II,” *Perspect. Biol. Med.* 46, no. 3 (Summer 2003): 383-402. On NIH's extramural research program, see Daniel Fox, “The Politics of the NIH Extramural Program, 1937-1950,” *J. Hist. Med. Allied Sci.* 42 (1987): 447-66; Stephen Strickland, *Politics, Science, and Dread Disease: A Short History of United States Medical Research Policy* (Cambridge: Harvard University Press, 1972). On federal oversight and NIH's peer review system, see Richard Mandel, *A Half Century of Peer Review, 1946-1996* (Bethesda, Md.: Division of Research Grants, NIH, 1996); Sejal Patel, “Methods and Management: NIH Administrators, Federal Oversight, and the Framingham Heart Study,” *Bull. Hist. Med.* 86, no. 1 (Spring 2012): 94-121; Stephen Strickland, *The Story of the NIH Grants Program* (New York: New York University Press, 1989). On NIH contributions to biomedical computing and biostatistics, see Joseph November, *Biomedical Computing: Digitizing Life in the United States* (Baltimore: Johns Hopkins University Press, 2012); Sejal Patel, “The Benevolent Tyranny of Biostatistics: Public Administration and the Promotion of Biostatistics at the National Institutes of Health, 1946-1970,” *Bull. Hist. Med.* 87, no. 4 (Winter 2013): 622-47. On NIH's contributions to research ethics, see Richard Mandel, *Beacon of Hope, 1953-1993: The Clinical Center Through Forty Years of Growth and Change in Biomedicine* (Bethesda, Maryland: National Institutes of Health, 1993); Laura Stark, *Behind Closed Doors: IRBs and the Making of Ethical Research* (Chicago: University of Chicago Press, 2012).

12 Sources were identified in residual files retrieved from the Office of NIH History Archives, National Institutes of Health, Bethesda, MD (hereafter ONH Archives); the Office of the Director, National Institutes of Health, Building 1, Bethesda, MD (hereafter Residual Files); and the National Cancer Institute's LION database, Bethesda, MD (hereafter NCI LION). For a comprehensive guide to NIH historical documents, see David Cantor, “Finding Historical Records at the National Institutes of Health,” *Soc. Hist. Med.* 28 (2015): 617-37.

Fig. 1. The frequency by year of NIH Consensus Development Program conferences in three periods. See <https://consensus.nih.gov/previous.htm> and <https://consensus.nih.gov/historical.htm>.

management, breast cancer screening, hyperactivity disorders, and estrogen therapy.¹³ However, none have examined the CDP itself as an object of historical analysis. Philosophers of science have written about the process of consensus in science more generally, exploring its limitations as a form of knowledge production and as an elusive goal.¹⁴ Most notably, Miriam Solomon has examined the history of the CDP as part of a broader investigation of scientific knowledge production in the United States since the late 1970s. Solomon argues that NIH consensus conferences were activities designed to lend credibility to existing consensus among researchers rather than establish new consensus or resolve scientific controversy.¹⁵ Leaving aside questions of the program's effectiveness, this article engages with the CDP as a means of better understanding the ongoing struggle of NIH officials to come to terms with the agency's mission as the major patron of biomedical research in the United States. Rather than

13 See Jeremy Greene, *Prescribing by Numbers: Drugs and the Definition of Disease* (Baltimore: Johns Hopkins University Press, 2007); Barron Lerner, *The Breast Cancer Wars: Fear, Hope, and the Pursuit for a Cure in Twentieth-Century America* (New York: Oxford University Press, 2001); Matthew Smith, *An Alternative History of Hyperactivity: Food Additives and the Feingold Diet* (New Brunswick: Rutgers University Press, 2011); and Elizabeth Siegel Watkins, *The Estrogen Elixir: A History of Hormone Replacement Therapy in America* (Baltimore: Johns Hopkins University Press, 2007).

14 Steve Fuller, *Social Epistemology* (Bloomington: Indiana University Press, 1991); David Guston, "Stabilizing the Boundary between US Politics and Science: The Role of the Office of Technology Transfer as a Boundary Organization," *Soc. Stud. Sci.* 29, no. 1 (Feb. 1999): 87-111.

15 Miriam Solomon, "Group Judgment and the Medical Consensus Conference," *Handbook Philos. Sci.* 16 (2010): 239-254; Miriam Solomon, *Making Medical Knowledge* (Oxford: Oxford University Press, 2015).

addressing all 159 conferences, this article relies on several representative case studies to explore the selection of topics, mechanisms of achieving consensus, standards of evidence, and the NIH's relationship with the medical and lay communities.

The exact nature of the NIH's mission was a focus of debate since the agency's founding. When the Ransdell Act formalized the National Institute of Health in 1930, it defined the NIH mission as encompassing the "study, investigation, and research in the fundamental problems of the diseases of man and matters pertaining thereto."¹⁶ Congressional reviews of NIH activities during Shannon's tenure, including hearings by Congressman Lawrence Fountain (D-NC) during the late 1950s, similarly focused on the goal of producing new biomedical knowledge, without mention of its dissemination to the profession or the public. In 1963, the NIH Study Committee, appointed by President Kennedy's Office of Science and Technology and commonly referred to as the Wooldridge Committee, reviewed the efficiency and accountability of NIH administrative structures and practices. In its 1965 report, the Committee reiterated the broad mandate of the agency, concluding that the NIH mission entailed "the stimulation and support of a very broad range of health-related or biomedical research."¹⁷ While it noted that the agency's aim was "improvement of virtually all aspects of the Nation's health," the Committee argued that this broader goal was attainable primarily with research in the basic sciences.¹⁸

In their internal, unpublished response to the Wooldridge report, NIH administrators drew particular attention to the question of mission and the existing tension between basic and applied research. Commentary prepared by the NIH Director's office posited that the committee's report "conveys a sense that the NIH is perceived as a science agency interested in biological and medical problems, and not a health agency applying science and technology to the solution of health and health-related problems."¹⁹ While the official stance emphasized the agency's research functions, many NIH administrators believed it was a "health agency," which ought to be more directly involved in the business of "the diagnosis, treatment, and prevention of disease."²⁰ This underlying tension reflected the essential debate that would characterize more than thirty years of the CDP. In theory, the CDP was designed to help NIH officials demonstrate the clinical relevance of the research they funded in Bethesda and at academic medical centers across the country. Examining its history illuminates how officials sought to define the agency's role at a time when calls for public accountability in science were becoming increasingly frequent. But how could the NIH respond to these calls without neglecting its mission as the nation's primary producer of biomedical knowledge? By gathering experts to discuss how new insights from biomedical research could be applied in the medical and public health arenas, NIH

16 Public Law 71-251, US Statutes at Large, 1171 (1930).

17 US NIH Study Committee, *Biomedical Science and Its Administration* (Washington, D.C.: Government Printing Office, 1965), 3.

18 *Ibid.*, 2.

19 "Biomedical Science and Its Administration. The Wooldridge Committee Report. The Initial NIH Commentary (April 1965)," 9, ONH Archives.

20 *Ibid.*, 13.

administrators assured that the agency was paying attention to the clinical utility of the research it funded.

NIH institutes employed the CDP for slightly different purposes at different times: as a response to perceived mission creep, as a defense against external criticism, and as an incubator for other NIH programming. The conferences highlighted in this article illustrate how NIH administrators hoped to use the program to engage in public discourse with physicians and other medical professionals (the CDP's original intended audience) and sometimes with the general public while recognizing that public dissatisfaction with medical experts was high.²¹ Between 1977 and 2013, the NIH grew from eleven institutes to twenty, some of which were the products of division or consolidation. Unsurprisingly, the NCI, the historically best-funded institute, was the most frequent sponsor or co-sponsor of CDP conferences. Other institutes, particularly NIA (Aging), NICHD (Child Health and Human Development), and NIDDK (Diabetes and Digestive and Kidney Diseases) were consistent conference sponsors, doing so during each decade of the CDP's existence. A review of CDP consensus statements suggests that all but one NIH institute sponsored or co-sponsored at least one CDP conference (Figure 2). As OMAR director Barnett Kramer later reflected, "we were supposed to be a resource to the other institutes."²²

OMAR carefully planned conferences to avoid the perception of undue influence by the sponsoring institute. This way, institutes "could legitimately say that they recognized a gap in knowledge. . .but they did not want to be accused of having forced the conclusion in one direction or another."²³ The mechanism of institute sponsorship was critical to maintaining the administrative balance of power among the Office of the Director, OMAR, and the institutes, demonstrating that CDP topics were selected by individual institutes, rather than by the central NIH governance. As Kramer explained, "Building 1 [where all central administration was located] wouldn't put forward ideas. And we [OMAR] were in Building 1, so they supported the process; they funded us. But they. . .didn't play an intimate role beyond that."²⁴

This paper focuses on three distinct periods in the CDP's history. These three periods, which I define as formation, codification, and retrenchment, were characterized by changing approaches to the project of consensus making. The CDP's first five years represented a transformative period of institutional formation (1977-1981) during which NIH administrators worked to define and publicize the CDP. A prolonged codification period (1982-1995) followed, during which NIH administrators held frequent consensus conferences aimed at influencing clinicians and policymakers, even while they continued to debate the agency's mission. Guided by a series of commissioned

21 Solomon suggests that "although the NIH has emphasized two purposes of its consensus conferences—to bring about consensus and disseminate it—the NIH Consensus Development Conference Program may have had other functions which are not mentioned (maybe not even noticed), or not mentioned as often" (*Making Medical Knowledge*, 60.) This paper identifies and illustrates some of those lesser-noticed functions.

22 Kramer, Barnett. Interview with the author. Rockville, Maryland, May 30, 2018.

23 Kramer, Barnett. Interview with the author. Rockville, Maryland, May 30, 2018.

24 Ibid.

Fig. 2. The frequency of sponsorship and co-sponsorship of NIH Consensus Development Program conferences by individual Institutes and Centers.

Abbreviations: NCI, National Cancer Institute; NICHHD, National Institute for Child Health and Development; NIAMDD, National Institute of Arthritis, Metabolism, and Digestive Diseases; NIADDKD, National Institute of Arthritis, Diabetes, and Digestive and Kidney Diseases; NHLBI, National Heart, Lung, and Blood Institute; NIA, National Institute on Aging; NINCDS, National Institute of Neurological and Communicative Disorders and Stroke; NINDS, National Institute of Neurological Disorders and Stroke; ADAMHA, Alcohol, Drug Abuse, and Mental Health Administration; NIMH, National Institute of Mental Health; NINR, National Institute of Nursing Research; NCNR, National Center for Nursing Research; NIAID, National Institute of Allergy and Infectious Diseases; NIAMS, National Institute for Arthritis and Musculoskeletal Diseases; NIDA, National Institute on Drug Abuse; CC, Clinical Center; OAM, Office of Alternative Medicine; NCCAM, National Center for Complementary and Alternative Medicine; NCHGR, National Center for Human Genome Research; NHGRI, National Human Genome Research Institute; NIAAA, National Institute on Alcoholism and Alcohol Abuse; NIDCD, National Institute on Deafness and Other Communication Disorders; NLM, National Library of Medicine; NIEHS, National Institute of Environmental Health Sciences; ODS, Office of Dietary Supplements; NEI, National Eye Institute; NIGMS, National Institute of General Medical Sciences; NCRR, National Center for Research Resources; ORMH, Office of Research on Minority Health; NIBIB, National Institute of Biomedical Imaging and Bioengineering

external program reviews in the 1980s and 1990s, the CDP underwent frequent reforms. Limited by a reticence to intrude upon physician autonomy or provide prescriptive advice, the CDP found itself unable to enforce its recommendations or even evaluate the extent to which the medical community was aware of them. Nevertheless, this period was marked by several tangible successes such as insurance coverage of liver transplantation, the adoption of newborn auditory screening, and the implementation of clinical trials registries.

This long period of continuity gave way to a period of retrenchment (1996-2013) that included a yearlong pause in conference activity during which the CDP underwent a systematic reevaluation that determined that the program needed to change how evidence was collected and disseminated in advance of conferences. In response, new OMAR leadership ushered in a period during which far fewer conferences were held.

This retrenchment period coincided with two notable developments that transformed science governance at the NIH: the broad institutionalization of translational medicine under the NIH directorships of Elias Zerhouni and Francis Collins, and a five-year doubling of the NIH budget soon followed by recession-related budgetary constraints. By 2013, the CDP was formally terminated, perceived by administrators as an increasingly irrelevant tool amid the dominance of multiple other organizations tasked with evaluating evidence and developing clinical practice guidelines.

CONSENSUS AND THE TECHNOLOGICAL IMPERATIVE

During the 1960s, discussions in Congress about linking research to practice emerged in considerations of technology assessment, a distinctive field of policy research that arose in the setting of concerns about the financial and social costs of new technologies.²⁵ The introduction of Medicare and Medicaid in 1965, which represented a significant new federal and state financial commitment to health care, only added to these concerns.²⁶ More generally, Americans worried that the cost of health care was spiraling out of control, spurred by President Nixon's 1969 declaration of a "massive crisis" in health care.²⁷

By the mid-1970s, the federal government had initiated several efforts to investigate and regulate the organization of research at the NIH. In 1975, Public Law 93-352 authorized the creation of the President's Biomedical Research Panel, a seven member body which would "review and assess the conduct, support, policies, and management of biomedical and behavioral research" at the NIH.²⁸ The panel's 1976 report concluded—just as President Kennedy's earlier Wooldridge Committee report had—that there remained an essential tension "between the research mission of the NIH and the application and dissemination of knowledge."²⁹ In fact, the committee found that an average of seven years elapsed between a scientific discovery and its introduction into clinical practice.³⁰

As the federal financial commitment to health care continued to grow, so too did Congressional inquiries. In June 1976, the Senate Labor and Public Welfare Subcommittee on Health convened to discuss one of the presumed culprits in rising health care costs: new clinical technologies such as coronary bypass surgery and kidney dialysis. Senators Jacob Javits (D-NY) and Edward Kennedy (D-MA) argued that the NIH had a responsibility to evaluate the cost-effectiveness of these new modalities.³¹ NIH director Donald Fredrickson agreed, remarking, "the NIH and the rest of the

25 H. David Banta and Clyde J. Behney, "Policy Formulation and Technology Assessment," *Milbank Mem Fund Quart Health Soc* 59 (1981): 445-79.

26 Jonathan Oberlander, *The Political Life of Medicare* (Chicago: University of Chicago Press, 2003); Laura Katz Olson, *The Politics of Medicaid* (New York: Columbia University Press, 2010).

27 Robert Hackey, *Cries of Crisis: Rethinking the Health Care Debate* (Reno: University of Nevada Press, 2012): 8.

28 United States Department of Health, Education, and Welfare, *Report of the President's Biomedical Research Panel* (Washington, D.C.: US Government Printing Office, 1976), v.

29 United States Department of Health, Education, and Welfare, *Report of the President's Biomedical Research Panel*, 5.

30 Ibid., Appendix B, 3-16.

31 See Edward Kennedy, "Where Should the Money Go," *The Sciences* 16 (Nov-Dec 1976): 10-11, 27.

scientific community must assume greater responsibility for the effect of research on the quality and cost of health care.”³² Fredrickson had already sensed the need for the NIH to defend its mission and productivity more broadly. One year earlier, he had argued that a “translation gap” from research findings to clinical applications hindered the agency’s efforts to provide directly useful knowledge to the American public.³³

In response to these concerns, Fredrickson commissioned NIH hematologist Seymour Perry to produce a report on how the NIH might translate biomedical research into clinical practice. Released in February 1977, the report emphasized the need for stronger collaboration between the NIH and other members of the medical community. It also recommended establishing “a consensus-building mechanism,” which would “include representatives of major participants in health care delivery, financing, and regulation.”³⁴ This new mechanism would provide a venue for evaluating ambiguous clinical evidence, particularly for new or unproven technologies. A new office formed within the Office of the Director would develop guidelines for identifying new clinically relevant research knowledge and disseminate that information to medical practitioners.³⁵

Perry strove to establish clearly delineated boundaries for the proposed consensus “mechanism.” As he acknowledged, professional societies and academic medical centers were already engaged in translating research into practice, albeit ineffectively in his estimation. The consensus “mechanism” would carry some advisory capacity, and greater collaboration among the federal government and medical professionals would represent a “final and essential step in the transfer of new information pertinent to health and disease, [which] is not being approached systematically.”³⁶ Perry estimated start-up costs totaling only \$669,000 for the first year, which included funding to hire five personnel for an office within the Office of the Director, embed ten personnel across five unspecified NIH institutes, develop an information dissemination program, and sponsor five consensus conferences at an approximate cost of \$35,000 each.³⁷

In response to Congressional inquiries, Frederickson requested in May 1977 that the Department of Health, Education, and Welfare establish an Office of Medical Applications of Research (OMAR) that would facilitate technology transfer activities across all NIH institutes. Its mandate was wide-ranging. As announced in the *Federal Register* in October 1978, OMAR’s responsibilities included advising the NIH director, monitoring technology transfer activities, and serving as the liaison to other health agencies and the broader health care community.³⁸ In other words, the new office

32 Testimony of Donald S. Fredrickson, M.D., Director, National Institutes of Health, before the Senate Labor and Public Welfare Subcommittee on Health, 17 June 1976.

33 Donald S. Fredrickson, “The Purposes of the National Institutes of Health: I. On the Translation Gap” (Bethesda, MD: National Institutes of Health, July 1975), ONH Archives.

34 “The Responsibilities of the National Institutes of Health at the Health Research/Health Care Interface” (Bethesda, MD: National Institutes of Health, Feb. 1977), 2, ONH Archives.

35 *Ibid.*, 9.

36 *Ibid.*, Appendix A, 7.

37 *Ibid.*, 11.

38 Public Health Service, National Institutes of Health, “Statement of Organization, Functions, and Delegations of Authority,” *Federal Register* 43, no. 199 (October 13, 1978): 47285.

would identify, collect, and disseminate clinically relevant knowledge to the broader medical community and public.

Fredrickson and Perry maintained that OMAR and the CDP would accomplish several administrative priorities at NIH. First, the new initiatives would bring health professionals and other stakeholders together to assess the most current, best evidence in the spirit of improving health care. Second, they would strengthen the NIH's role as a disseminator of biomedical information while still preserving its core mission of producing biomedical research. But NIH administrators debated how the agency should engage in public debates regarding new technologies. Should it privilege technical factors or evaluate social, ethical, or economic ones? This was in some ways a false dichotomy. Indeed, the interconnectedness of social, ethical, economic, and technical considerations accounted, in part, for the subsequent tensions that later emerged regarding the CDP's activities.

One of Fredrickson's first challenges as NIH director included navigating heated public debates about the safety and ethics of recombinant DNA technology and deciding whether and how NIH should create guidelines for research in this area following the introduction of Paul Berg's pioneering, albeit controversial, gene-splicing technique.³⁹ The NIH ultimately issued guidelines for recombinant DNA research in June 1976, around the same time Fredrickson was testifying before Congress about rising health care costs. His experience with the recombinant DNA controversy highlighted a question that would challenge future OMAR administrators: what role, if any, should the NIH play in addressing social concerns about new research technologies?⁴⁰ Fredrickson and fellow NIH administrators determined that the CDP would assess the scientific and technical aspects of new clinical technologies, but not necessarily their economic, ethical, and social implications. As Fredrickson explained, "if we lay out the state-of-the-art—what it is we know and do not know from data scientifically derived—we will serve medicine and society through provision of a sounder base on which further value judgments can be laid."⁴¹ But avoiding the social and ethical implications of these technologies would prove difficult and sometimes impossible.

INSTITUTIONAL FORMATION: 1977-1981

OMAR sponsored thirty-one conferences during the CDP's first five years, most between 1978 and 1980. These first five years were characterized by multiple attempts to define the CDP both institutionally and intellectually.⁴² OMAR staff faced the burden of fitting the concept of consensus development into the existing structure of biomedical research while also distinguishing it as a new enterprise. This forced them to reckon

39 See Sheldon Krimsky, *Genetic Alchemy: The Social History of the Recombinant DNA Controversy* (Cambridge: The MIT Press, 1982).

40 Donald S. Fredrickson, *The Recombinant DNA Controversy: A Memoir — Science, Politics, and the Public Interest; 1974–1981* (Washington, D.C., ASM Press, 2001).

41 Donald Fredrickson, "Seeking Technical Consensus on Medical Interventions," *Clin. Res.* 26 (1978): 116-7, 117.

42 Seymour Perry, "The NIH Consensus Development Program: A Decade Later," *JAMA* 317, no. 8 (Aug. 20, 1987): 485-88.

with the essential role of the NIH. Was it to be a patron of knowledge production, an information clearinghouse, or some combination of the two? Administrators at other federal agencies, most notably the National Science Foundation (NSF), were also grappling with how to define their institutional missions. The contemporaneous rise of the NSF and NIH after World War II would have a lasting impact on postwar scientific research in the United States. As Toby Appel has demonstrated, NIH's rise as a major patron of biomedical research challenged NSF's identity as a sponsor of basic research, and in response, NSF administrators sought to distinguish its patronage efforts from those of the NIH by focusing primarily on the production of new knowledge, with minimal attention to its clinical application.⁴³

As OMAR staff members established the CDP's intellectual and institutional foundations, they frequently described its mission as though repetition would solidify its purpose. Early CDP pamphlets framed the CDP as "aimed at improving the translation of the results of biomedical research. . . into knowledge that can effectively be employed in the practice of medicine and public health."⁴⁴ Yet, officials alternately described the CDP as a communication program, a peer review program, and a technology assessment program. One early OMAR administrator made the case for the second interpretation, arguing that the peer review system and the CDP differed only on a matter of scale. In one sense, he argued that the CDP was to be the peer review system for the entire biomedical community—effectively yet another venue for expert opinion.⁴⁵ Meanwhile, promotional materials described the CDP as a science communications tool. One information sheet explained that the CDP would "provide the physician and the public with current, responsible information on the pros and cons of medical technologies."⁴⁶ Panelists at one 1979 CDP conference, for example, discussed the benefits and risks of wisdom tooth extraction, while participants of a 1980 conference on febrile seizures discussed the role and potential risks of seizure prophylaxis.⁴⁷ In this sense, the CDP functioned as a kind of medical advisor of sorts.

The general contours of the CDP's approach to conference planning and administration also took shape during this first period. Typically, NIH institute directors approached OMAR and requested that it plan and administer conferences on topics of particular significance to their institute. These requests dictated CDP conference topics, evident by the consistent variety of topics addressed in any given year.⁴⁸ A

43 Toby Appel, *Shaping Biology: The National Science Foundation and American Biological Research, 1945-1975* (Baltimore: Johns Hopkins University Press, 2000).

44 "OMAR Information Sheet," Office of Medical Applications of Research, National Institutes of Health (April 1980), NCI LION.

45 Kalberer, "Peer Review," *Sci. Technol. Human Values* 10, no. 3 (Summer 1985): 63-72.

46 "NIH Consensus Development," Office of Medical Applications of Research, National Institutes of Health, n.d. (ca. 1980), ONH Archives.

47 "Removal of Third Molars," *NIH Consens. Statement Online* 1979 Nov 28-30; 2(11): 65-68; "Febrile Seizures," *NIH Consens. Statement Online* 1980 May 19-21; 3(2): 1-10.

48 Nor were conference topics directly related to the professional interests of OMAR directors and acting directors, who included a cardiovascular epidemiology specialist (William Friedewald), former FDA administrator (J. Richard Crout), hematologist (Seymour Perry), neurologist (John Ferguson), oncologist (Barnett Kramer), pediatrician (Charles Lowe), and engineer without medical training (Itzhak Jacoby).

planning committee of OMAR staff and other NIH institute representatives then selected panelists and speakers and prepared the conference schedule.⁴⁹

Conferences typically lasted two or three days and consisted of a combination of open public sessions and closed executive sessions held on the NIH campus in Bethesda, Maryland. During the public sessions, speakers presented data and commentary to the conference panelists and audience members. Upon conclusion of the public sessions, panelists participated in an overnight executive session during which they crafted a preliminary CDP statement. On the final day of the conference, panelists distributed the statement to audience members and received feedback during one final round of public discussion. After this final plenary session, panelists and OMAR staff edited and finalized the statement, which then stood as the final record of the conference.

As OMAR staff codified CDP practices during the mid-1980s, CDP statements developed a consistent formula that required responses to between four and six questions. Panelists typically assessed the current state of knowledge regarding the chosen conference topic and in some cases offered practice or policy recommendations. After 1982, all CDP statements concluded with panelists' identification of priorities for future research. This longstanding practice represented the CDP as an endeavor operating at the intersection of basic and applied science. CDP research recommendations typically ranged from calls for further investigation into the natural history and mechanisms of particular diseases to outcomes evaluations of discrete diagnostic and therapeutic interventions.

Perhaps the most important part of a consensus conference happened after it ended. OMAR had amassed a broad network of contacts to which it sent CDP statements, including national TV networks, local news outlets, and an array of medical and scientific journals. Administrators touted this network when they referred to the CDP as a communications tool. Indeed, a 1989 RAND review of the CDP framed it as "more than an assessment program. It is also a communication program to the professional community and the public."⁵⁰ CDP conferences brought scientists, clinicians, and health advocates to the NIH campus and allowed them to utilize OMAR's public communication and media distribution expertise to spread their message. OMAR administrators hoped that their unique form of communication would facilitate behavioral changes among health care professionals. As OMAR director John Ferguson noted when he was appointed to the position in 1988, patient care obligations often limited his opportunities to learn about new medical technologies. The CDP, he explained, "brings the best,

49 Panelists were typically either professionals with established reputations or individuals known by OMAR administrators and CDP planning committee members. Some presumably received NIH grants, but at least from the mid-1990s onward, panelists could not have direct intellectual or financial conflicts of interest with the conference topic. For an early version of CDP planning guidelines, see Office of Medical Applications of Research, *Guidelines for the Selection and Management of Consensus Development Conferences* (Bethesda, MD: National Institutes of Health, 1983), ONH Archives.

50 David E. Kanouse et al., *Changing Medical Practice through Technology Assessment: An Evaluation of the National Institutes of Health Consensus Development Program* (Santa Monica, CA: The RAND Corporation, 1989): v.

most current science directly to the clinician. It summarizes the information and critiques it in as non-biased a manner as possible.”⁵¹ By the mid-1980s, the OMAR mailing list totaled over 21,000 recipients.⁵² Ideally, this would make the most state of the art medical information easily accessible to practicing clinicians.

Whether NIH administrators defined the CDP’s mission in terms of peer review or science communication, they wanted it to provide an interactive venue for bridging research and practice. A CDP conference would bring together multiple stakeholders to evaluate publicly an ambiguous, incomplete, or emerging evidence base on a particular medical technology. As an informational brochure explained, the CDP would “[improve] the translation of the results of biomedical research pertinent to health care into knowledge that can effectively be employed in the practice of medicine and public health.”⁵³ The language of this brochure suggests that even from the beginning, NIH officials believed that medical knowledge was too vast to be interpreted by individual practicing doctors and needed “translating” for public consumption.

The first CDP conference predated the formal organization of OMAR. Held over a three-day period as an open forum on the NIH campus, this September 1977 conference on breast cancer screening posed a series of questions to a panel composed of medical and public health professionals, professors of law and theology, and other members of the public. Only three of the sixteen panelists were women, including New York Medical College oncologist Jane Wright, who was later celebrated for her achievements as a pioneering African American female physician, and Yale Divinity School ethicist Margaret Farley. The conference panel was charged with evaluating evidence related to breast cancer screening, particularly data from the Breast Cancer Detection Demonstration Project.

The National Cancer Institute (NCI) and the American Cancer Society had jointly organized the Demonstration Project in 1974 to determine the role of different screening modalities in the early detection of breast cancer. In an era when feminists were attacking medical experts for their paternalistic attitudes and demanding more bodily autonomy, the Demonstration Project—and more broadly the issue of breast cancer screening—represented an opportunity for women to speak up and demand that they be more involved in their health care decisions.⁵⁴

By 1976, science journalists began to raise questions about the possible risks of the radiation from mammograms administered during the Demonstration Project. They were not alone. As the use of radiation for diagnostic and therapeutic procedures increased during the 1950s and 1960s, so too did public concerns about patient safety

51 “Ferguson to Head Office of Medical Applications of Research,” *NIH Record* (August 23, 1988): 8, https://nihrecord.nih.gov/PDF_Archive/1988%20PDFs/19880823.pdf.

52 Fitzhugh Mullan and Itzhak Jacoby, “The Town Meeting for Technology: The Maturation of Consensus Conferences,” *JAMA* 254, no. 8 (Aug. 23/30, 1985): 1068-72. See also, Itzhak Jacoby and Susan Clark, “Direct Mailing as a Means of Disseminating NIH Consensus Statements: A Comparison with Current Techniques,” *JAMA* 255 (1986): 1328-30.

53 “OMAR Information Sheet.”

54 Lerner, *The Breast Cancer Wars*.

and the hazards associated with radiation exposure.⁵⁵ Critics suggested that radiation exposure itself could cause future breast cancer in previously asymptomatic women, and the popular press quickly picked up on these findings. *San Francisco Chronicle* journalist Judith Randal called the Demonstration Project “a vicious circle of mistakes” while the *New York Times* dismissed the program as “ill-conceived” and “dangerous.”⁵⁶

In January 1977, the NCI had convened a review panel to determine the utility of routine breast cancer screening for asymptomatic women. The panel occurred at a time when breast cancer activists were publicly challenging the medical *status quo* regarding radical mastectomy, but before much was known about the science and implementation of population-based breast cancer screening.⁵⁷ It found no justification for routine mammographic screening for women under the age of fifty, recommending it only for women in their forties with a family history of breast cancer and for women between thirty-five and thirty-nine with a personal history of breast cancer.⁵⁸ The September 1977 CDP conference took place not long after the review panel released its report. In his opening remarks, NIH Director Donald Fredrickson expressed hope that the conference would hasten the “conversion of information to knowledge, and knowledge to wisdom.”⁵⁹ After two days of presentations, the CDP panel produced a statement that its members “deplored the lack of clear-cut data on the efficacy and the risk-benefit ratio of screening for women under 50 years of age,” and argued for new randomized clinical trials.⁶⁰ The CDP panel endorsed the recommendations of the earlier Demonstration Project review panel, effectively upholding the findings of an NIH-funded controversial program that faced serious criticisms from the general public.

The CDP panel did, however, recommend that Demonstration Project administrators develop clearer, more transparent informed consent documents and background materials for women and their physicians.⁶¹ Although these suggestions were not technical in nature, conference participants stressed their importance for encouraging better communication between physicians and patients. In his closing remarks, Fredrickson praised consensus development as “a method for speeding up the transfer and validation of a body of medical knowledge which can be applied for the common good.”⁶² Samuel

55 See Angela N.H. Creager, *Life Atomic: A History of Radioisotopes in Science and Medicine* (Chicago: University of Chicago Press, 2013): 311-350.

56 Judith Randal, “Mammography—A Vicious Circle of Mistakes,” *San Francisco Chronicle*, Oct. 6, 1977, 35; Daniel Greenberg and Judith Randal, “The Questionable Breast X-Ray Program,” *Washington Post*, May 1, 1977, C5; “Mammography Mistake,” *New York Times*, Sept. 25, 1977, 172.

57 On the history of breast cancer and breast cancer activism, see Lerner, *The Breast Cancer Wars*; Ellen Leopold, *A Darker Ribbon: Breast Cancer, Women, and Their Doctors in the Twentieth Century* (Boston: Beacon Press, 1999).

58 “Supplemental and Concluding Report of the Working Group to Review the National Cancer Institute-American Cancer Society Breast Cancer Detection Demonstration Projects,” *J. Natl. Cancer Inst.* 62, no. 3 (1979): 699-709.

59 Donald Fredrickson, “Opening Remarks: NIH/NCI Consensus Development Meeting on Breast Cancer Screening,” September 14, 1977, NCI LION.

60 “Breast Cancer Screening,” *NIH Consens. Statement Online* 1977 Sep 14-16; 1(1): 5-8.

61 *Ibid.*

62 Donald Fredrickson, “NIH Closing Statement: NIH/NCI Consensus Development Meeting on Breast Cancer Screening,” September 16, 1977, NCI LION.

Thier, the chair of internal medicine at Yale University School of Medicine and chair of the conference, published an editorial in the *New England Journal of Medicine* that accompanied the CDP statement. As he explained, the “breast-cancer screening controversy illustrates the risk of transferring technology too rapidly even when the transfer is made with the best of intentions.”⁶³ He was suggesting there simply was not sufficient data to make a clear recommendation about mammography. Had the consensus conference been premature?

The gap between the promise of a new technology and its most effective clinical use was made clear during another early CDP conference on cervical cancer screening in 1980. Many physicians made it a routine practice in the 1970s to recommend yearly Pap smears, while critics like the feminist news journal *Off Our Backs* commented that the practice “[provides] doctors with a steady, reliable basic income” and “explains why obstetrician-gynecologists are the wealthiest of all the medical specialists.”⁶⁴ In addition, by the late 1970s, the practice had come into question through the accumulating evidence that it took over a decade for a precancerous in-situ cervical lesion to develop into invasive cancer.⁶⁵

Reflecting on these findings, as well as the fact that that most of the data on Pap effectiveness was based on far more infrequent screening, the American Cancer Society (ACS) amended its recommendations that year to allow for screenings every three years in women who had had two negative Pap smears one year apart.⁶⁶ The American College of Obstetricians and Gynecologists (ACOG) maintained the importance of yearly Pap smears as a means of closely following women at higher risk for cervical cancer as well as providing better primary health care to all women.⁶⁷ Despite these debates, the benefit of cervical cancer screening remained unclear. Although screening advocates had maintained that Pap smears reduced mortality rates for cervical cancer, subsequent analyses suggested that screening had a smaller than expected impact.⁶⁸ Two epidemiologists went so far as to proclaim the Pap smear a “dubious policy success” because policymakers endorsed and health professionals performed the procedure before clear evidence of its efficacy was available.⁶⁹

In 1980, OMAR convened a CDP conference to review the existing evidence on screening intervals and cervical cancer mortality rates. After working until four in the

63 Samuel Thier, “Breast-Cancer Screening: A View from Outside the Controversy,” *N. Engl. J. Med.* 297, no. 19 (Nov. 10, 1977): 1063-65, 1065.

64 Fran Moira, “Intercourse & Cervical Cancer,” *Off Our Backs* 10, no. 8 (September 30, 1980): 4.

65 David Guzik, “Efficacy of Screening for Cervical Cancer: A Review,” *Am. J. Public Health* 68, no. 2 (1978): 125-34.

66 “Cancer of the Cervix,” *CA Cancer J. Clin.* 30, no. 4 (July/August 1980): 215-223.

67 Phil Gunby, “Aftermath of the ACS Bombshell,” *JAMA* 243, no. 23 (June 20, 1980): 2373-2374.

68 A.B. Miller, J. Lindsay, and G.B. Hill, “Mortality from Cancer of the Uterus in Canada and its Relationship to Screening for Cancer of the Cervix,” *Int. J. Cancer* 17 (1976): 600-12. On the delay between the invention and implementation of a medical technology, see Joel D. Howell, *Technology in the Hospital: Transforming Patient Care in the Early Twentieth Century* (Baltimore: Johns Hopkins University Press, 1995).

69 Anne-Marie Foltz and Jennifer Kelsey, “The Annual Pap Test: A Dubious Policy Success,” *Milbank Mem. Fund Quart. Health Soc.* 56 (Autumn 1978): 426-62.

morning to come to a consensus, the panel endorsed the Pap smear as a routine screening procedure, but recommended against screening for women who had never been sexually active and for those over the age of sixty who had two previous negative Pap smear results. Appeasing both the ACS and ACOG, the panel also concluded that healthy, low-risk women should be rescreened every one to three years, with the precise interval at the discretion of a woman and her physician.⁷⁰ In response, the feminist National Women's Health Network released a statement recommending screening every three years after three consecutive negative annual Pap smears, focusing on the issue of access to screening for disadvantaged women.⁷¹ Although it failed to please everyone, the split panel had issued a clear set of clinical guidelines, going far beyond mere evaluation of technology transfer or scientific advising. And yet the resulting CDP statement also strove to preserve patient autonomy by answering to women who sought control over their bodies with its flexible statement about screening intervals.

From its inception, the CDP fielded criticism from detractors who warned that consensus development would lead to potential government intrusion upon medical decision-making. In 1980 the director of the American College of Surgeons deemed the CDP an attention-getting maneuver by the NIH, concluding, "I find it difficult to see how this entire 'consensus' meddling can advance the best interests of the individual patient."⁷² Concerns such as this reflected decades-long physician fears of government control, which had been particularly apparent during recurrent debates about creating a national health system.⁷³

These criticisms were also the result of the CDP's nebulous character in the 1970s: an evaluative program that appeared to engage in quasi-regulatory activities. In fact, NIH administrators envisioned the CDP taking on an advisory role, not a regulatory one. Seymour Perry argued in an internal document distributed to institute directors that the NIH catalyzed the consensus-making enterprise.⁷⁴ Perry's language was purposeful; a catalyst is a substance that speeds up a chemical reaction without itself being altered. Thus, he maintained that the NIH was solely serving as a dispassionate third party advisor, putting the right people in the same room to evaluate and disseminate biomedical knowledge. Yet some critics argued that this very practice lacked a certain degree of analytical precision despite its NIH imprimatur. One commentator in the *New England Journal of Medicine* worried that CDP statements would prompt undue fear or halt innovation. As he remarked, "I fear lest these statements will act like a dead hand, discouraging thought. After all, in the nature of things, they can only rehash, codify, and grade work already done."⁷⁵ Indeed, consensus development relied heavily

70 "Cervical Cancer Screening: The Pap Smear," *NIH Consens. Statement Online* 1980 Jul 23-25; 3(4):1-11.

71 Moira, "Intercourse & Cervical Cancer," 4.

72 C. Rollins Hanlon, "Director's Memo, Consensus Development in Medicine," *Bull. Am. Coll. Surg.* 65 (August 1980): 1, 3, 3.

73 Paul Starr, *The Social Transformation of American Medicine: The Rise of a Sovereign Profession and the Making of a Vast Industry* (New York: Basic Books, 1982).

74 Seymour Perry, "Consensus Development Definition Paper" (Bethesda, MD: National Institutes of Health, November 14, 1978), NCI LION. See also Seymour Perry, "NIH Panels Analyze Issues on Technology," *Hospitals* 53 (June 1, 1979): 115-16, 115.

75 Drummond Rennie, "Consensus Statements," *N. Engl. J. Med.* 304 (March 12, 1981): 665-66, 666.

upon the role of expert opinion, as CDP panelists and presenters deliberated over the veracity and utility of extant knowledge. Yet in the eyes of NIH administrators, CDP conferences could help defend the agency against external criticism or promote internal programming, both of which had been a goal of the first conference on breast cancer screening.

Some conferences were held in response to public health concerns. Starting in February 1980, newspapers began reporting higher rates of Reye syndrome, a life threatening condition associated with swelling in the brain and fatty deposits in the liver that seemed to occur in children recovering from chickenpox or influenza.⁷⁶ In July, the Centers for Disease Control (CDC) reported on a possible link between aspirin and Reye syndrome based on a small case control study of seven children with the condition in Arizona.⁷⁷ In November, the CDC reported on similar studies from Ohio and Michigan, also suggesting an association between aspirin and Reye. Although the studies were small and imperfect, relying on subjects to remember accurately whether or not they had taken aspirin, the CDC concluded that aspirin “may be a factor in the pathogenesis of Reye syndrome” and urged parents to “use caution” when giving it to their ill children.⁷⁸

When the CDP conference on Reye syndrome was convened in March 1981, anxieties were high.⁷⁹ Tellingly, OMAR staff narrowed the conference focus to include only the diagnosis and treatment of Reye syndrome, not the emerging controversy surrounding its association with aspirin. This kept with the NIH’s stated preference for technical consensus but also enabled NIH officials to avoid wading into the growing public debate. Ultimately, the CDP statement made only brief reference to the aspirin association. Noting that “salicylates [such as aspirin] alone cannot be responsible for the development of Reye’s syndrome,” the statement stressed the “need for further carefully designed studies before recommending changes in antipyretic therapy of children.”⁸⁰

The statement was ahead of its time. A year later, the CDC, Surgeon General, and American Academy of Pediatrics all issued definitive statements declaring an association between aspirin use and Reye syndrome. In 1986, the FDA ruled that all aspirin bottles were required to carry a warning about the Reye syndrome link.⁸¹ In a sense,

76 E.R. Shipp, “15 in State Struck by Reye’s Syndrome since Jan. 1,” *New York Times*, March 2, 1980, 21.

77 Follow-up on Reye Syndrome — United States,” *MMWR* 29, no. 27 (July 11, 1980): 321-22.

78 “Reye Syndrome — Ohio, Michigan,” *MMWR Morb. Mortal. Wkly. Rep.* 29, no. 44 (November 7, 1980): 532, 537-39, 539.

79 “The Diagnosis and Treatment of Reye’s Syndrome,” *NIH Consens. Statement Online* 1981 Mar 2-4; 4(1): 1-15. See also, Karen M. Starko, C. George Ray, Lee B. Dominguez, Warren L. Stromberg, and Dora F. Woodall, “Reye’s Syndrome and Salicylate Use,” *Pediatrics* 66, no. 6 (1980): 859-64.

80 “The Diagnosis and Treatment of Reye’s Syndrome.”

81 Committee on Infectious Diseases, “Aspirin and Reye Syndrome,” *Pediatrics* 69 (1982): 810-12, 811; “Surgeon General’s Advisory on the Use of Salicylates and Reye Syndrome,” *MMWR Morb. Mortal. Wkly. Rep.* 31, no. 22 (June 11, 1982): 289-90; Stephen B. Soumerai, Dennis Ross-Degnan, and Jessica Spira Kahn, “Effects of Professional and Media Warnings about the Association between Aspirin Use in Children and Reye’s Syndrome,” *Milbank Quart.* 70 (1992): 155-82, see 164-65; Associated Press, “Aspirin Labels to Warn about Reye Syndrome,” *New York Times*, March 8, 1986, 7. On the history of Reye syndrome and aspirin, see Mark A. Largent, *Keep Out of Reach of Children: Reye’s Syndrome, Aspirin, and the Politics of Public Health* (New York: Bellevue Literary Press, 2015) and Cynthia Connolly, *Children and Drug Safety: Balancing Risk and Protection in Twentieth-Century America* (New Brunswick: Rutgers University Press, 2018).

the Reye conference had preserved institutional autonomy for the NIH. The agency had responded to public pressure by hosting the Reye conference, but the decision to sidestep controversy ensured that other health agencies—in this case, the CDC and FDA—would be more actively engaged in media coverage of the protracted controversy while the NIH maintained its focus on clinical research.

CODIFICATION: 1982–1995

In 1982, OMAR administrators found themselves pushed to refine the CDP's identity in light of the newly created National Center for Health Care Technology, a new subunit within the US Department of Health and Human Services that was completely independent of NIH.⁸² Although the National Center was short-lived (it existed only from 1979 to 1982), it served as a counterpoint to NIH's approach to technology assessment. The Health Services Research, Health Statistics, and Health Care Technology Act of 1978 gave the National Center a congressional mandate to assess the safety and efficacy of health care technologies, instructing the Center to develop and distribute norms and standards for the use of health care technologies. The Act also gave the Center a quasi-regulatory role, advising the Health Care Financing Administration on Medicare reimbursement policies.⁸³

In contrast, the CDP generally limited its focus to the safety and efficacy of new or unproven technologies. This narrow scope reflected the NIH preference to focus on the technical rather than the social or economic aspects of new treatments and technologies.⁸⁴ The prioritization of "technical consensus" and subtle avoidance of the social or cultural aspects of medical advances ensured that the NIH's interpretation of technology assessment remained closely tied to its primary mission as a biomedical research agency. When the Reagan administration imposed major budget cuts in 1982, the National Center was dissolved, while OMAR escaped unscathed, in part because it did not identify as a policymaking body like the National Center, an easy target for the anti-regulatory administration.⁸⁵

Both former and acting OMAR officials continued to grapple with how the CDP's activities pertained to the NIH mission. Former OMAR director Charles Lowe and his assistant Susan Asch noted in 1984 that "In simplest terms, OMAR is intended to be the 'facilitator' that will permit basic and, in particular, clinical research to be transferred more rapidly into health care practice."⁸⁶ Lowe and Asch maintained that the CDP did

82 For more on the National Center for Health Care Technology, see David Blumenthal, "Federal Policy Toward Health Care Technology: The Case of the National Center," *Milbank Mem. Fund Q. Health Soc.* 61 (1983): 584-613; Seymour Perry, "The Brief Life of the National Center for Health Care Technology," *N. Engl. J. Med.* 307 (1982): 1095-1100.

83 Health Services Research, Health Statistics, and Health Care Technology Act of 1978, Public Law 95-623, US Statutes at Large, 92 (1978).

84 "The Responsibilities of the National Institutes of Health," ONH Archives; Perry, "Consensus Development Definition Paper;" Paul M. Wortman, Amiram Vinokur, and Lee Sechrest, "Do Consensus Conferences Work? A Process Evaluation of the NIH Consensus Development Program," *J. Health Polit. Policy Law* 13 (1988): 469-98.

85 Blumenthal, "Federal Policy."

86 Susan Asch and Charles Lowe, "The Consensus Development Program: Theory, Process, and Critique," *Knowledge S* (March 1984): 369-85; 376.

not attempt to encourage practitioners to adopt consensus conference findings in the clinic, as this would exceed the NIH's domain of expertise and authority.⁸⁷ The issue of what kinds of questions OMAR ought to address also provoked some handwringing. In particular, administrators wondered whether CDP conferences ought to address social, ethical, legal, or economic questions, worrying that such topics were beyond their mandate.⁸⁸ Ultimately, they agreed to focus only on new biomedical knowledge, theoretically stripped of its sociocultural factors. Although the CDP was unable to resolve questions of authority and mission completely, it hosted eight-seven conferences between 1982 and 1995, averaging at least six conferences per year. This codification period represented the most active time for the CDP in terms of conference frequency, program evaluation, and public engagement.

Despite these mission-related challenges, several CDP conferences during this period spurred meaningful changes in health policy.⁸⁹ In 1983, OMAR co-sponsored a conference on liver transplantation in order to evaluate which patients would be optimal candidates to receive an organ and to identify the long-range risks and benefits of transplantation. With the introduction of the immunosuppressant drug cyclosporine in 1980, liver transplants had become more feasible and far more common. But cyclosporine cost \$6000 per year and was not covered by Medicare, and liver transplants were expensive interventions in their own right.⁹⁰ The decision to hold a consensus conference on liver transplantation was, according to OMAR's deputy director John Kalberer, "a political request" derived from public pressure. As OMAR associate director Richard Crout recalled, "The topic was picked basically by HCFA [Health Care Finance Administration] facing the decision on whether to reimburse" insurance companies for the costs of the operation and the care associated with it.⁹¹

In the CDP statement, the panel concluded that although more research was necessary, "liver transplantation offers an alternative therapeutic approach which may prolong life in some patients suffering from severe liver disease that has progressed beyond the reach of currently available treatment."⁹² The CDP conference played a critical role in the acceptance and coverage of transplantation as a standard therapeutic option. For example, a 1984 Illinois statute stated that transplantation of any organ would only be covered by insurance if OMAR deemed it to be a non-experimental procedure.⁹³ Thus a CDP statement effectively

87 Ibid., 379.

88 John Ferguson, "NIH Consensus Conferences: Dissemination and Impact," *Ann. NY Acad. Sci.* 703 (1993): 180-199; 184.

89 Mae Thamer, Nancy Fox Ray, Scott Henderson, Cheryl Rinehart, Charles Sherman and John Ferguson, "Influence of the NIH Consensus Conference on *Helicobacter Pylori* on Physician Prescribing among a Medicaid Population," *Med. Care* 36, no. 5 (May 1998): 646-60.

90 Gerald Markle and Daryl Chubin, "Consensus Development in Biomedicine: The Liver Transplant Controversy," *Milbank Quart.* 65, no. 1 (1987): 1-24, 6-7.

91 Ibid., 9.

92 "Liver Transplantation," *NIH Consens. Statement Online* 1983 Jun 20-23; 4(7): 1-15.

93 Markle and Chubin, "Consensus Development in Biomedicine." On the procedure's broader acceptance, see Roger W. Evans, Diane L. Manninen, and Frederick B. Dong, "An Economic Analysis of Liver Transplantation," *Gastroenterol. Clin. N. Amer.* 22, no. 2 (1993): 451-73.

served as a guiding document for the policymaking process—even as NIH officials argued that the CDP statements were not supposed to have any regulatory capacity.

A 1988 CDP conference on the cochlear implant was met with mixed public response. Although cochlear implants had been developed in the 1960s and implanted in several thousand individuals by the time of the conference, questions about the success and safety of the devices remained unresolved. At the time of the CDP conference, the FDA classified cochlear implants for children as investigational devices because of a lack of data on how implants influenced long-term auditory development. The conference thus focused on identifying suitable candidates for implantation, assessing benefits and risks for different types of cochlear implants, and evaluating special considerations for implantation in children. The resulting CDP statement, which deemed cochlear implants “an important step in our long-range goal of understanding, preventing, and treating hearing impairment and resulting language disorders,” was favorable but noncommittal.⁹⁴ It reflected support for the assistive devices while calling for additional research on auditory pathways, comparative analyses of different device models, and new assessment and education programs.

Members of the deaf community were more circumspect. Many deaf adults and parents of deaf children argued that the cochlear impact represented a tacit rejection of Deaf identity and culture.⁹⁵ Yet the CDP statement’s favorable assessment of cochlear implants gained traction in the medical community across the Atlantic. In the United Kingdom, the CDP statement played a critical role in the funding of these devices by the British National Health Service, a rare but important example of CDP activities influencing global health practices.⁹⁶ The CDP similarly played a role in the adoption of newborn auditory screening. A 1993 conference held on the topic recommended universal auditory screening of all newborns, after which state officials from Rhode Island and California cited the CDP statement when establishing their own state-based universal screening programs.⁹⁷

Around the same time, several CDP conferences and workshops addressed the dissemination of clinical trial data. In the early 1990s, researchers, policymakers,

94 “Cochlear Implants,” *NIH Consens. Statement Online* 1988 May 2-4; 7(2): 1-25.

95 See Stuart Blume, *The Artificial Ear: Cochlear Implants and the Culture of Deafness* (New Brunswick, NJ: Rutgers University Press, 2010). Tension within the deaf community remained unresolved a decade after the conference: see Marie Arana Ward, “As Technology Advances, a Bitter Debate Divides the Deaf,” *The Washington Post*, May 11, 1997, A1. On the history of Deaf culture, see John V. Van Cleve, ed., *The Deaf History Reader* (Washington, DC: Gallaudet University Press, 2007).

96 Ibid. For a broader discussion of technology assessment and the cochlear implant, see Pascale Lehoux and Stuart Blume, “Technology Assessment and the Sociopolitics of Health Technologies,” *J. Health Polit. Policy Law* 25, no. 6 (December 2000): 1083-1120.

97 Betty R. Vohr, Lisa M. Carty, Patricia E. Moore, and Kristen Letourneau, “The Rhode Island Hearing Assessment Program: Experience with Statewide Hearing Screening (1993-1996),” *J. Pediatr.* 133 (September 1998): 353-57; Senate Health and Human Services Committee Analysis, SB 555, February 24, 1997, http://www.leginfo.ca.gov/pub/97-98/bill/sen/sb_0551-0600/sb_555_cfa_19970516_155242_sen_comm.html.

and health advocates debated the merits of democratizing access to clinical trial data.⁹⁸ A 1991 CDP workshop discussed how and when early, practice-changing results of NIH-funded clinical trials should be communicated to the public, a process workshop participants felt was clinically and morally important.⁹⁹ A second workshop in 1993 considered the issue of clinical trial registries. While two individual NIH institutes had their own online databases with information about ongoing trials, there existed no comparable, comprehensive database for all government-funded clinical trials.¹⁰⁰ The workshop statement concluded that clinical trial registries were an integral component of an “evidence based health system” and emphasized the need to capture both published and unpublished data so that medical professionals would have access to all relevant data in making clinical decisions.¹⁰¹ By 1997 The Food and Drug Administration Modernization Act required the NIH to create a public registry of clinical trials and in 2000, the NIH developed ClinicalTrials.gov, which remains the most comprehensive clearinghouse for clinical trials data.

Nevertheless, NIH administrators struggled throughout the 1980s and 1990s to convince a hesitant medical community that the CDP would not intrude upon the professional autonomy of clinical practitioners. As one physician warned in a *JAMA* editorial, “we must protect the individual choices of each physician from the potential tyrannical domination of consensus and allow the process of development of new knowledge to continue.”¹⁰² To assuage these concerns, administrators continually stressed the CDP’s advisory nature.¹⁰³ It was almost as though OMAR doth protest too much – as though continuing to declare the office’s role would obscure the uncertainty that remained about its real function. But CDP statements on topics like Pap smears, which had included prescriptive recommendations that resembled clinical guidelines in language and style, gave the appearance that the CDP engaged in the very practices that NIH officials were trying to avoid.¹⁰⁴

98 Ian Chalmers, “Underreporting Research is Scientific Misconduct,” *JAMA* 263 (1990): 1405-8; Henry Greenberg, “Access to Data: A Contemporary Direction for Clinical Trials,” *Control Clin Trials* 13 (1992): 93-6; RD Morris, J Lau, NJ Arena, FE Nardine, and TC Chalmers, “A Clinical Trials Database as a Research Tool in Health Care,” *Online J. Curr. Clin. Trials* (17 Jul 1992), Doc No 14; AJ Munro, “Publishing the Findings of Clinical Research,” *BMJ* 308 (1993): 1340-1.

99 Clinical Trials Results: Exploring the Dissemination Process, <https://consensus.nih.gov/1991/1991ClinicalTrialsDisseminationta008html.htm>.

100 These included the NCI’s Physician Data Query and National Institute of Allergy and Infectious Diseases’s AIDSTRIALS. See SM Hubbard, NB Martin, LW Blankenbaker, RJ Esterhay Jr, DR Masys, MC Stram, and VT DeVita Jr, “The Physician Data Query (PDQ) Cancer Information System,” *J. Cancer Educ.* 1 (1986): 79-87; RJ Veenstra and JC Gluck, “Access to Information about AIDS,” *Ann. Intern. Med.* 114 (1991): 320-4.

101 An Evidence Based Health Care System: The Case for Clinical Trials Registries, <https://consensus.nih.gov/1993/1993EvidenceBasedTrialRegistriesta013html.htm>.

102 William E. May, “Consensus or Coercion,” *JAMA* 254 (Aug. 23/30, 1985): 1077.

103 John Ferguson, “Technology Transfer: Consensus and Participation. The NIH Consensus Development Program,” *Jt. Comm. J. Qual. Improv.* 21 (July 1995): 332-36, 333.

104 On the history of clinical practice guidelines, see George Weisz, Alberto Cambrosio, Peter Keating, Loes Knappen, Thomas Schlich, and Virginia J. Tournay, “The Emergence of Clinical Practice Guidelines,” *Milbank Quart.* 85 (Dec. 2007): 691-727.

From the beginning, NIH administrators had maintained that the CDP was designed to facilitate changes in physician behavior. However, without the authority to implement or enforce its recommendations, the program had great difficulty instituting change. Moreover, it proved difficult to demonstrate whether the CDP had accomplished such change, or even whether physicians were aware of CDP statements despite OMAR's robust distribution network.¹⁰⁵ NIH administrators attributed these challenges to what they deemed the "two-cultures" problem. CDP conference organizers were typically members of the academic community and representatives of the federal government, but practicing physicians and other clinicians were the intended audience.¹⁰⁶ Perhaps the communities were failing to communicate. But in attempting to avoid the perception of intruding upon the professional autonomy of medical professionals, NIH officials had left the CDP devoid of any teeth to implement its recommendations.

Despite lacking any enforcement or legal regulatory power, the CDP was nagged throughout the 1980s and 1990s by accusations that it had engaged in favoritism by prioritizing particular views over others. In one such case, one 1984 CDP panel had recommended that all Americans over the age of two reduce their consumption of cholesterol and saturated fats in order to prevent coronary artery disease based on evidence from a clinical trial of a cholesterol-lowering drug. Critics questioned the strength of evidence employed by the CDP panel, arguing that data was not yet strong enough to warrant changing dietary recommendations for an entire population.¹⁰⁷ A 1990 conference on the use of recombinant bovine somatotropin (rBST) on dairy farms to increase milk production concluded that the hormone had no measurable effect on the milk produced by treated cows.¹⁰⁸ Although OMAR officials hoped the conference would resolve an ongoing public controversy, the panelists—including a dairy farmer—had not had access to unpublished data from rBST manufacturers. Critics concluded that the NIH was merely in the pocket of the dairy industry. As one physician opponent sniped, "Essentially, the panel has examined sanitized data of industry scientists and their indentured academics."¹⁰⁹

Hoping to respond to some of these critiques, NIH administrators commissioned numerous internal and external evaluations of the CDP during this codification period.¹¹⁰ These evaluations focused on both procedural changes, such as offering continuing medical education credits for conference participants, and on more substantive

105 Jacqueline Kosecoff, David Kanouse, William Rogers, Lois McCloskey, Constance Monroe Winslow, and Robert H. Brook, "Effects of the National Institutes of Health Consensus Development Program on Physician Practice," *JAMA* 258 (1987): 2708-13.

106 Ferguson, "NIH Consensus Conferences," 189.

107 Gina Kolata, "Heart Panel's Conclusions Questioned," *Science* 227 (January 4, 1985): 40-41; Michael Oliver, "Consensus or Nonsense Conferences on Coronary Artery Disease," *Lancet* 1 (1985): 1087-1089, 1088.

108 "Bovine somatotropin," *NIH Technol. Assess. Statement Online* 1990 Dec 5-7; (7):16.

109 Ann Gibbons, "Bovine Hormone Gets the Nod," *Science* 250 (Dec. 14, 1990): 1506.

110 See Jacqueline Kosecoff et al., "Effects;" Institute of Medicine, *Consensus Development at the NIH: Improving the Program* (Washington, D.C., National Academy Press, 1990).

matters of ethics and tone.¹¹¹ In response, OMAR officials strengthened conflict of interest guidelines by barring panelists from having a financial interest in a conference topic and by requiring planning committee members to disclose both intellectual and financial conflicts of interest.¹¹² The result, they hoped, would be panels comprised of more neutral participants who could sift through the available evidence more impartially. As Miriam Solomon has argued, these procedural changes bolstered the “appearance of trustability” associated with consensus conference proceedings.¹¹³

RETRENCHMENT (1996-2013): THE COLLAPSE OF CONSENSUS

By the late 1990s, CDP activities entered a downturn. OMAR had sponsored an average of at least six conferences a year during the preceding two decades, a number which was cut in half between 1996 and 1999. Yet consensus conferences continued to address complex issues and skirt the boundary between knowledge dissemination and guideline creation. Conferences on total knee replacement (2003) and hepatitis B (2008) drew explicit attention to perceived shortcomings in evidence-based practices, while the hepatitis B conference and another on acupuncture (1997) also engaged with broader programming priorities for NIH institutes.

A 1997 conference on breast cancer screening emerged as perhaps the most controversial consensus conference. In January 1997, the CDP convened yet another conference on breast cancer screening.¹¹⁴ Four years earlier, the NCI had stopped recommending mammograms for women in their forties based on the findings of its International Workshop on Screening for Breast Cancer, a decision which sparked public controversy and Congressional inquiry.¹¹⁵ In response, NCI Director Richard Klausner requested that a CDP conference address the controversy and evaluate new data. The conference became notable for the firestorm that ensued upon its adjournment. The CDP statement endorsed the NCI’s decision, stating that available scientific data did not support a recommendation for universal mammography screening for

111 Itzhak Jacoby, “The Consensus Development Program of the National Institutes of Health. Current Practices and Historical Perspectives,” *Int. J. Technol. Assess. Health Care* 1 (April 1985): 419-432.

112 “Guidelines for the Planning and Management of NIH Consensus Development Conferences,” (Bethesda, MD, National Institutes of Health, May 1993; updated March 1995), NCI LION. See also John Ferguson, “The NIH Consensus Development Program. The Evolution of Guidelines,” *Int. J. Technol. Assess. Health Care* 12(1996): 460-474.

113 Solomon, *Making Medical Knowledge*, 57.

114 For more on this conference and broader debates about breast cancer screening, see Robert Aronowitz, *Unnatural History: Breast Cancer and American Society* (Cambridge: Cambridge University Press, 2007); David Cantor, “Cancer Screening,” in Mark Jackson, ed., *The Routledge History of Disease* (London: Routledge, 2017): 358-77; Lerner, *Breast Cancer Wars*; and Ilana Löwy, *Preventive Strikes: Women, Precancer, and Prophylactic Surgery* (Baltimore: Johns Hopkins University Press, 2010).

115 Suzanne W. Fletcher et al., “Report of the International Workshop on Screening for Breast Cancer,” *J. Natl. Cancer Inst.* 85 (1993): 1644-56; House Committee on Government Operations, *Misused Science: The National Cancer Institute’s Elimination of Mammography Guidelines for Women in Their Forties*. Union Calendar No 480. House Report 103-863. October 20, 1994 (Washington, D.C.: United States Government Printing Office, 1994).

women in their forties.¹¹⁶ As Constance Rufenbarger, a consumer advocate, breast cancer survivor, and CDP panelist, best summed up: “There’s no one up here who would not like to say that if you go for that mammogram, we can find early cancer and you would be cured. But if you cannot say that, can you recommend to every woman in her 40’s that she have a mammogram?”¹¹⁷ Many felt this was an oversimplification of the matter, however.

The CDP statement angered screening proponents, many of whom argued that the panelists misinterpreted data, overemphasized the risks of screening, and underplayed its benefits.¹¹⁸ One New Mexico mammographer and conference speaker called the statement “tantamount to a death sentence for thousands of women.”¹¹⁹ Other journalists noted that “other than the muffin crumbs from its meetings, the only thing this panel left behind is more confusion,” and that “the wimpy panel could not bring itself to recommend the most effective diagnostic test we have.”¹²⁰ Senator Arlen Specter, then chairman of the Labor, HHS, and Education appropriations committee, called for a Congressional hearing. Senator Olympia Snowe, along with a bipartisan group of fifty-one co-sponsors, ultimately introduced a non-binding “sense of the Senate” resolution that urged the NCI to consider recommending mammography screening for women in their forties or to “direct the public to consider guidelines issued by other organizations.”¹²¹ The resolution passed by a 98-0 vote. As unique as the firestorm might have been, the conference had done what it was meant to do: evaluate ambiguous scientific information related to medical practice.¹²²

Although this conference highlighted the NIH’s often difficult position within a broader debate about clinical guidelines, prevention, and women’s health, it is important to recognize that CDP conferences that ended so divisively were not the norm. While conceived in response to criticism of cancer screening practices, the CDP statement opened the door for public criticism of the NIH, particularly when the statement was at odds with guidelines issued by other professional organizations. Although NIH administrators made it abundantly clear that OMAR was not in the business of producing clinical guidelines and had adopted neutralizing practices like choosing panel members without conflicts of interest, some CDP activities and statements gave the

116 “Breast Cancer Screening for Women Ages 40-49,” *NIH Consens. Statement Online* 1997 Jan 21-23; 15(1): 1-35.

117 Gina Kolata, “Mammogram Talks Prove Indefinite,” *New York Times* (Jan. 24, 1997): A1.

118 Daniel Kopans, “The Breast Cancer Screening Controversy and the National Institutes of Health Consensus Development Conference on Breast Cancer Screening for Women Ages 40-49,” *Radiology* 210 (1999): 4-9.

119 Kolata, “Mammogram Talks,” A1.

120 “Cancer Copout,” *Newark Star-Ledger*, January 26, 1997, 2; Laura Pulfer, “Breast Cancer a Deadly Test of Health Care,” *Cincinnati Enquirer*, January 28, 1997, 3.

121 143 *Cong. Rec.* S970 (1997).

122 In fact, NCI Director Richard Klausner questioned the CDP panel’s conclusions during a press conference, remarking, “I’m giving you my own interpretation, and yes I am disagreeing with the balance [of evidence presented].” See “Press Conference: Consensus Development Conference on Breast Cancer Screening for Women Ages 40-49 Held January 23, 1997,” NCI LION.

appearance that the NIH was participating in the very activities its leaders insisted fell outside their purview.

A 1997 conference on acupuncture marked the CDP's entry into yet another long-standing debate within clinical research and medicine: how to evaluate practices deemed as alternative or outside the dominant biomedical paradigm. The 1991 creation of the NIH's Office of Alternative Medicine had proven contentious as critics were dismayed by the prospect of federal funds being spent to evaluate alternative medicine while the nation was experiencing a recession.¹²³ Six NIH institutes served as conference co-sponsors, suggesting relatively wide support for NIH evaluation of alternative medical practices.

CDP panelists were charged with reviewing data on the efficacy and clinical applications of acupuncture as well as identifying barriers to its broader adoption. As the panel maintained, research on acupuncture's efficacy consisted primarily of case studies rather than controlled trials. Even so, the panel acknowledged that "[t]here is sufficient evidence of acupuncture's value to expand its use into conventional medicine and to encourage further studies of its physiology and clinical value."¹²⁴ Soon after the statement's publication, critics expressed frustration. Stanford oncologist Wallace Sampson, founder of the *Scientific Review of Alternative Medicine*, argued that conference panelists "engaged in pseudoscientific reasoning" and produced a report that suffered from a "lack of critical, scientific thinking."¹²⁵ In *Science*, Yale anesthesiologist Arthur Taub derided the conference as a meaningless farce. "In short," he concluded, "it appears that the panel concluded that acupuncture was virtually useless, declared a 'victory' as ordered up, and called for more research expenditure to heap on that already wasted."¹²⁶ With charges that the CDP was serving a political, rather than a scientific function, NIH Director Harold Varmus decided it was time to reevaluate the program.

In October 1998, Varmus established a working group to evaluate CDP practices and procedures, consisting of five NIH administrators and four outside professionals.¹²⁷ The working group concluded that the CDP still had an important role to play "in informing health care professionals and the public [sic] about improvements in medical practice and public health, and in clarifying inconsistencies between practice and supporting evidence."¹²⁸ Several changes were recommended, including a "more formalized and rigorous" process for selecting conference topics, preliminary meetings to discuss the evidence in advance of the actual conference, and the commissioning of evidence reviews to be distributed to panelists ahead of time.¹²⁹ This reassessment

123 Eric Boyle, "The Politics of Alternative Medicine at the National Institutes of Health," *Federal History Journal* 3 (2011): 16-32.

124 "Acupuncture," *NIH Consens. Statement Online* 1997 Nov 3-5; 15(5):1-34.

125 Wallace Sampson, "Critique of the NIH Consensus Conference on Acupuncture," <https://www.acu-watch.org/general/nihcritique.shtml>.

126 Arthur Taub, "Thumbs Down on Acupuncture," *Science* 279 (January 9, 1998): 159.

127 Members included Alan Leshner, Ezra Davidson, Peggy Eastman, Scott Grundy, Barry Kramer, Joan McGowan, Audrey Penn, Stephen Straus, and Steven Woolf.

128 "Report of the Working Group of the Advisory Committee to the Director to Review the Office of Medical Applications of Research" (June 1999), 10, Residual Files, 2.

129 Ibid.

coincided with a complete halt in activity for the CDP. That year, OMAR sponsored no CDP conferences and John Ferguson retired after eleven years as OMAR director.¹³⁰

With a new decade came realignment for the CDP. In 2000, NCI deputy director Barnett Kramer was appointed as Ferguson's successor. Kramer had served as an institute representative for the 1997 breast cancer conference and was also a member of the 1999 Working Group.¹³¹ Kramer's tenure as OMAR director formalized two initiatives begun under Ferguson's directorship: the institution of a "firewall" between sponsoring NIH institutes and OMAR, and the implementation of systematic evidence reviews preceding all conferences. As Kramer explained, "we wanted to make sure that the firewall was thick and strong so we developed a formal process that would take us to the point that the institute would recede into the background and leave the actual process of the conference to us."¹³² A "firewall" could thus deflect claims of institutes' undue influence on consensus statements, which marred some consensus conferences during the 1980s.

In 2001, OMAR established a partnership with the Agency for Healthcare Research and Quality (AHRQ), another agency within the US Department of Health and Human Services, in order to produce and deliver systematic evidence reviews to conference panelists about one month before each CDP conference, as recommended by the Working Group.¹³³ This new arrangement extended the CDP's long-standing practice of sending conference panelists a packet that included conference abstracts and other background materials.¹³⁴ It also signified the CDP's incorporation of evidence-based methodologies as a means of further bolstering its claim to relevance. This move reflected OMAR administrators' awareness of the growing influence of evidence-based medicine, albeit at least a decade after which the concept gained popularity.¹³⁵ And yet, by providing more rigorous background information to participants in the form of systematic reviews, OMAR administrators were demonstrating their embrace of evidence-based medicine while, as Miriam Solomon has argued, potentially obviating the need to hold conferences at all.¹³⁶

The decreased annual frequency of conferences continued for the remainder of the CDP's tenure. In fact, during eight of the CDP's final fourteen years, the CDP sponsored only one or two conferences per year. A 2003 conference on total knee replacement highlighted how different groups of clinical practitioners perceived the strength of evidence (or lack thereof) supporting longstanding clinical practices, thus positioning OMAR uncomfortably between competing professional interests. The conference

130 Carla Garnett, "OMAR's Longest-Serving Director Ferguson Retires," *NIH Record*, September 21, 1999, https://nihrecord.nih.gov/newsletters/09_21_99/retirees.htm.

131 Kramer, Barnett. Interview with the author. Rockville, Maryland, May 30, 2018.

132 Ibid.

133 "Report of the Working Group," 10.

134 Mullan and Jacoby, "The Town Meeting for Technology," 1071.

135 Stefan Timmermans and Marc Berg, *The Gold Standard: The Challenge of Evidence-Based Medicine* (Philadelphia: Temple University Press, 2003).

136 Solomon, *Making Medical Knowledge*, 51.

panel, which acknowledged the safety and cost-effectiveness of total knee replacement, nevertheless questioned the clinical benefits derived from the procedure. As Kramer explained, “there was a sense from some surgeons that they saw benefits with their own eyes and they were uncomfortable questioning whether replacing a knee was of benefit or not.”¹³⁷ Moreover, when the panelists also concluded that no data supported specific rehabilitative practices, “the physical therapists were very upset that there was a statement to the effect that we don’t know the nature of benefit—if it exists with physical therapy after the procedure.”¹³⁸ No conference was free of politics, although OMAR remained uneasy with this fact.

Although CDP conferences became less frequent affairs during this time, NIH institutes continued to use them to advance different programmatic goals. The 1997 acupuncture conference illustrated how multiple NIH institutes indicated general interest in evaluating alternative medical practices. And although a 2008 hepatitis B (HBV) conference was “an eye opener for some of the hepatologists” when it illustrated a disconnect between commonly used clinical endpoints and actual health outcomes,¹³⁹ it also represented a promising example of programming coordination, conveniently falling just after the announcement of a new collaborative research initiative sponsored by the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK).¹⁴⁰

With the appointment of radiologist and biomedical engineer Elias Zerhouni as NIH director in 2002 came a new focus on improving the quality of translational research. In 2003, Zerhouni launched the NIH-wide Roadmap for Medical Research, which formalized the agency’s commitment to convert research findings into new clinical applications.¹⁴¹ These presumably forward-thinking priorities actually mirrored those expressed in OMAR promotional materials from the early 1980s.¹⁴² Three decades later, NIH administrators were still debating the agency’s role within the broader clinical research enterprise.¹⁴³

In 2009, OMAR’s budget of \$4.713 million represented a two-fold inflation-adjusted increase since 1977, far less than the three-fold increase in overall NIH appropriations over that same span.¹⁴⁴ With the onset of the recession, 2009 also marked a new period of budgetary constraints for the NIH. The American Recovery and

137 Kramer, Barnett. Interview with the author. Rockville, Maryland, May 30, 2018.

138 Ibid.

139 Ibid.

140 Michael F. Sorrell et al., “National Institutes of Health Consensus Development Conference Statement: Management of Hepatitis B,” *Ann. Intern. Med.* 150 (2009): 104-10; Edward Doo and Jay Hoofnagle, “The Hepatitis B Research Network,” *Hepatology* 49, 5 Suppl (May 2009): S196-S197.

141 Elias Zerhouni, “The NIH Roadmap,” *Science* 302 (2003): 63-64; 72.

142 “OMAR Information Sheet.”

143 Solomon argues that translational medicine emerged as a “predictable corrective to the methodological limitations” of evidence-based medicine. See Solomon, *Making Medical Knowledge*, 13.

144 National Institutes of Health, FY 2010 Budget Congressional Justification Narrative Revised, Office of the Director, April 27, 2009, https://officeofbudget.od.nih.gov/pdfs/FY10/FY%202010%20CJ%20narrative%20revised%204-27-09_9pdf.pdf.

Reinvestment Act of February 2009 provided only \$10 billion to the NIH over a two-year period, and administrators across the NIH tightened their purse strings.¹⁴⁵

Meanwhile, the push for translational medicine at the NIH only intensified with the appointment of geneticist Francis Collins as NIH director. In 2010, the NIH filed a trademark application with the US Patent and Trademark Office for a new motto, “NIH. . . Turning Discovery Into Health,” granted in 2012, which captured the agency’s intention to emphasize and publicize its translational obligations.¹⁴⁶ In 2012, Collins established the National Center for Advancing Translational Sciences (NCATS) in order to build new partnerships with the private sector and more quickly develop clinical applications of basic research. As he remarked, “in recent years, researchers have succeeded in identifying the causes of nearly 4,500 diseases, but we have been unable to turn this knowledge into many new therapies: effective treatments exist for only about 250 of these diseases.”¹⁴⁷ In its first fiscal year, NCATS received \$574.7 million in funding.¹⁴⁸ This push for improved translational science ignored the CDP, which had been engaged in translating biomedical discoveries for over three decades. In comparison, the 2011 budget request prepared by the Office of the Director set OMAR’s estimated budget at a paltry \$4.935 million.¹⁴⁹

In January 2012, all OMAR staff and resources were transferred to the Office of Disease Prevention (ODP), another subunit of the Office of the Director. The announcement was buried five pages into an issue of the *NIH Record*, the NIH’s biweekly employee newsletter, and published almost six weeks after the fact. NIH director Francis Collins explained that the reorganization would confer “considerable operational efficiencies” and that CDP conferences would be held even less frequently.¹⁵⁰ In other words, NIH administrators perceived the CDP as an expendable program but still saw it as a valuable tool in high-profile cases. In 2013, the CDP was officially disbanded. In a perfunctory statement, the ODP acknowledged that other organizations—the Cochrane Collaboration, the Community Preventive Services Task Force, the Institute of Medicine, and the United States Preventive Services Task Force, among others—had assumed authoritative roles in conducting literature reviews, building consensus, and making guidelines, rendering the CDP redundant.¹⁵¹ Although Kramer had

145 National Institutes of Health, FY 2018 Budget, Office of the Director https://officeofbudget.od.nih.gov/pdfs/FY18/insti_center_subs/OD.pdf.

146 US Department of Health and Human Services. NIH: Turning Discovery Into Health. Registration Number 4168707. Filed December 6, 2010, and registered July 3, 2012.

147 Testimony Before the Health Subcommittee of the Energy and Commerce Committee United States House of Representatives, Statement of Francis S. Collins, M.D., Ph.D., Director, National Institutes of Health, US Department of Health and Human Services, June 21, 2012, <http://energycommerce.house.gov/sites/republicans.energycommerce.house.gov/files/Hearings/Health/20120621/HHRG-112-IF14-Wstate-CollinsF-20120621.pdf>.

148 National Center for Advancing Translational Sciences, Past Budgets, <https://ncats.nih.gov/about/center/budget/past>.

149 National Institutes of Health, FY 2011 Budget, Office of the Director, <https://officeofbudget.od.nih.gov/pdfs/FY11/FY%202011%20OD%20CJ%20508%20compliant.pdf>.

150 “OMAR Is Assumed Into ODP,” *NIH Record*, February 17, 2012, 5, https://nihrecord.nih.gov/pdfs/2012/04172012_Record.pdf.

151 “Consensus Development Program,” <http://consensus.nih.gov/>.

returned to the NCI by that time, he perceived that OMAR's closure was the result of the agency's sensibility that "we didn't want to do something another agency was doing now," particularly during a period of budget constraints.¹⁵²

BRIDGING THE TRANSLATION GAP

Over the course of several decades, politicians and NIH officials grappled with whether to define the NIH as a science agency focused on medical problems or a health agency that applied science to solve health-related problems. Charged with bridging the gap between research and practice, OMAR and the CDP—always small, circumscribed entities within the NIH bureaucracy—represented proxies for this broader debate about the NIH mission. Established in response to external calls for increased accountability of government funding, OMAR and the CDP allowed NIH administrators to test the waters with respect to publicizing and disseminating the clinical applications of basic research. Administrators presented the CDP as an information clearinghouse on new medical technologies while also affirming its neutrality and separation from regulatory and policymaking mechanisms. But they also used the CDP to respond to the political pressure they felt to defend and support agency-funded research while protecting the autonomy of their own member institutes. In essence, they hoped the CDP might serve as a buffer, protecting the NIH from larger cultural criticisms about the value and expense of research and new technologies.

By continually emphasizing information dissemination as a key driver of change, the CDP seemed to be operating under to a quixotic mandate, to change medical practice while eschewing regulation. For NIH administrators, the CDP was an opportunity to manage the pace of technological change by providing a new venue through which both established and emerging technologies could be evaluated. In some cases, the findings of the CDP directly led to changes in clinical practice and insurance coverage of new technologies and procedures, such as the 1983 liver transplant conference. In other cases, the findings of the CDP sparked stronger discord, such as the 1997 conference on mammography. And in many instances, the CDP had little or no influence at all.

If the work of the CDP was less about resolving scientific controversy, it was more about determining how the NIH could advance the clinical applications of the basic research it funded in Bethesda and across the country. In one sense, OMAR represented a boundary organization, operating at the intersection of research and clinical practice.¹⁵³ Not unlike technology transfer efforts in which incentives (such as licensing and patent protection) to commercialize scientific innovation are paramount, OMAR and the CDP provided opportunities for NIH administrators and researchers to demonstrate a certain degree of public accountability for NIH-supported research.

Almost forty years after the founding of the CDP, bridging the "gap" between research and practice remained an important goal for NIH administrators and researchers. For example, NIH's congressional justification for the 2013 fiscal year proclaimed

152 Kramer, Barnett. Interview with the author. Rockville, Maryland, May 30, 2018.

153 David Guston, "Stabilizing the Boundary Between US Politics and Science: The Role of the Office of Technology Transfer as a Boundary Organization," *Soc. Stud. Sci.* 29, no. 1 (Feb. 1999): 87-111.

that while fifty-four percent of the overall NIH budget had been devoted to basic research, the rest of the budget supported applied research.¹⁵⁴ The fact that the NIH supported basic and applied research in relatively equal measure demonstrates how NIH administrators had yet to resolve fully a question posed fifty years prior: was the NIH a science agency devoted to addressing medical concerns, or was it a health agency that applied science and technology in the name of improving health?

Ultimately, other initiatives superseded OMAR and the CDP during the NIH-wide turn toward translational medicine. As the NIH devoted more resources toward new initiatives such as the NCATS in an era of constrained budgets, OMAR remained relatively small. There is some irony in how NIH efforts to leverage evidence in pursuit of better health care fueled debate by providing a new arena in which professional conflicts could play out. Ultimately, the CDP failed to resolve essential tensions between science and medicine, research and practice, and information dissemination and medical regulation. But for over thirty-five years, it provided a critical space for debating not just important issues in health and biomedical research, but the mission and identity of the NIH as the nation's primary medical research agency.

ACKNOWLEDGEMENTS

Many thanks to Justin Barr, Deborah Doroshow, Robert Hackey, Chin Jou, and Naomi Rogers for reading and commenting on this article; David Cantor and the late Robert Martensen for their early advice; Barnett Kramer and the late John Ferguson for conversations about OMAR's institutional history; Richard Mandel for providing access to the residual files of the NIH Office of the Director; and the two anonymous reviewers who provided constructive feedback.

FUNDING

DeWitt Stetten, Jr., Fellowship in the History of Biomedical Sciences and Technology, Office of NIH History; School of Professional Studies Summer Scholarship, Providence College.

154 National Institutes of Health, FY 2013 Budget Congressional Justification, Office of the Director, https://officeofbudget.od.nih.gov/pdfs/FY13/FY2013_Overview.pdf.

Appendix A: Complete List of NIH Consensus Conferences

* designates OMAR workshop, Technology Assessment Workshop, or State-of-the-Science Conference

FORMATION (1977–1981)

1977

- Breast Cancer Screening

1978

- Educational Needs of Physicians and Public Regarding Asbestos Exposure
- Dental Implants: Benefit and Risk
- Mass Screening for Colorectal Cancer
- Treatable Brain Diseases in the Elderly
- Indications for Tonsillectomy and Adenoidectomy: Phase 1
- Availability of Insect Sting Kits to Non-Physicians
- Mass Screening for Lung Cancer
- Supportive Therapy in Burn Care
- Surgical Treatment of Morbid Obesity

1979

- Pain, Discomfort, and Humanitarian Care
- Antenatal Diagnosis
- Transfusion Therapy in Pregnant Sickle Cell Disease Patients
- Improving Clinical and Consumer Use of Blood Pressure Measuring Devices
- Treatment of Primary Breast Cancer: Management of Local Disease
- Steroid Receptors in Breast Cancer
- Intraocular Lens Implantation
- Estrogen Use and Postmenopausal Women
- Amantadine: Does It Have a Role in the Prevention and Treatment of Influenza
- The Use of Microprocessor-Based “Intelligent” Machines in Patient Care
- Removal of Third Molars

1980

- Thrombolytic Therapy in Thrombosis
- Febrile Seizures
- Adjuvant Chemotherapy of Breast Cancer
- Cervical Cancer Screening: The Pap Smear
- Endoscopy in Upper GI Bleeding
- Cesarean Childbirth
- CEA as a Cancer Marker
- Coronary Artery Bypass Surgery: Scientific and Clinical Aspects

1981

- The Diagnosis and Treatment of Reye's Syndrome
- Computed Tomographic Scanning of the Brain

CODIFICATION (1982–1995)

1982

- Defined Diets and Childhood Hyperactivity
- Total Hip Joint Replacement
- Clinical Applications of Biomaterials

1983

- Critical Care Medicine
- Liver Transplantation
- Evaluating the Elderly Patient: The Case for Assessment Technology*
- Treatment of Hypertriglyceridemia
- Precursors to Malignant Melanoma
- Drugs and Insomnia: The Use of Medications to Promote Sleep
- Dental Sealants in the Prevention of Tooth Decay

1984

- Diagnostic Ultrasound Imaging in Pregnancy
- Analgesic-Associated Kidney Disease
- Osteoporosis
- Mood Disorders: Pharmacologic Prevention of Recurrence
- Fresh Frozen Plasma: Indications and Risks
- Limb-sparing Treatment of Adult Soft-tissue Sarcomas and Osteosarcomas
- Lowering Blood Cholesterol to Prevent Heart Disease

1985

- Travelers' Diarrhea
- Health Implications of Obesity
- Anesthesia and Sedation in the Dental Office
- Donor Registries for Bone Marrow Transplantation*
- Electroconvulsive Therapy
- Adjuvant Chemotherapy for Breast Cancer

1986

- Health Implications of Smokeless Tobacco Use
- Prevention of Venous Thrombosis and Pulmonary Embolism
- Integrated Approach to the Management of Pain

- The Utility of Therapeutic Plasmapheresis for Neurological Disorders
- Impact of Routine HTLV-III Antibody Testing of Blood and Plasma Donors on the Health of the Public
- Infantile Apnea and Home Monitoring
- Platelet Transfusion Therapy
- Diet and Exercise in Noninsulin-Dependent Diabetes Mellitus

1987

- Newborn Screening for Sickle Cell Disease and Other Hemoglobinopathies
- Management of Clinically Localized Prostate Cancer
- Differential Diagnosis of Dementing Diseases
- Neurofibromatosis
- The Health Benefits of Pets*
- Geriatric Assessment Methods for Clinical Decision Making
- Magnetic Resonance Imaging

1988

- Prevention and Treatment of Kidney Stones
- Cochlear Implants
- Dental Implants
- Perioperative Red Cell Transfusion
- Urinary Incontinence

1989

- Therapeutic Endoscopy and Bleeding Ulcers
- Oral Complications of Cancer Therapies: Diagnosis, Prevention, and Treatment
- Modeling in Biomedical Research: An Assessment of Current and Potential Approaches: Applications to Studies in Cardiovascular/Pulmonary Function and Diabetes*
- Sunlight, Ultraviolet Radiation, and the Skin
- Treatment of Destructive Behaviors in Persons With Developmental Disabilities

1990

- Noise and Hearing Loss
- Surgery for Epilepsy
- Treatment of Sleep Disorders of Older People
- Adjuvant Therapy for Patients With Colon and Rectum Cancer
- Intravenous Immunoglobulin: Prevention and Treatment of Disease
- Health Care Delivery Research Using Hospital Firms*
- Report on the Workshop on Diffusion of ECMO Technology*
- Treatment of Early Stage Breast Cancer

- Diagnosis and Management of Asymptomatic Primary Hyperparathyroidism
- Clinical Use of Botulinum Toxin
- Bovine Somatotropin*

1991

- Clinical Trials Results: Exploring the Dissemination Process*
- Gastrointestinal Surgery for Severe Obesity
- Effects and Side Effects of Dental Restorative Materials*
- Treatment of Panic Disorder
- Diagnosis and Treatment of Depression in Late Life
- Acoustic Neuroma

1992

- Diagnosis and Treatment of Early Melanoma
- Triglyceride, High Density Lipoprotein, and Coronary Heart Disease
- Methods for Voluntary Weight Loss and Control*
- The Head and Heart of Chaos: Nonlinear Dynamics in Biological Systems*
- Gallstones and Laparoscopic Cholecystectomy
- Impotence

1993

- Early Identification of Hearing Impairment in Infants and Young Children
- Mortality and Morbidity of Dialysis
- Ultrasound Screening: Implications of the Radius Study*
- An Evidence-Based Health Care System: The Case for Clinical Trials Registries*

1994

- *Helicobacter Pylori* in Peptic Ulcer Disease
- Effect of Antenatal Steroids for Fetal Maturation on Perinatal Outcomes
- Ovarian Cancer: Screening, Treatment, and Followup
- The Persian Gulf Experience and Health*
- Optimal Calcium Intake
- Total Hip Replacement
- Bioelectrical Impedance Analysis in Body Composition Measurement*

1995

- Infectious Disease Testing for Blood Transfusions
- Gaucher Disease: Current Issues in Diagnosis and Treatment*
- Cochlear Implants in Adults and Children

- Integration of Behavioral and Relaxation Approaches Into the Treatment of Chronic Pain and Insomnia*
- Physical Activity and Cardiovascular Health

RETRENCHMENT (1996–2013)

1996

- Cervical Cancer
- Management of Temporomandibular Disorders*

1997

- Breast Cancer Screening for Women Ages 40-49
- Interventions to Prevent HIV Risk Behaviors
- Management of Hepatitis C
- Genetic Testing for Cystic Fibrosis
- Acupuncture
- Effective Medical Treatment of Opiate Addiction

1998

- Rehabilitation of Persons With Traumatic Brain Injury
- Diagnosis and Treatment of Attention Deficit Hyperactivity Disorder

1999

No Conferences Held In 1999

2000

- Improving Medical Implant Performance Through Retrieval Information: Challenges and Opportunities*
- Osteoporosis Prevention, Diagnosis, and Therapy
- Antenatal Corticosteroids Revisited: Repeat Courses
- Phenylketonuria (PKU): Screening and Management
- Adjuvant Therapy for Breast Cancer

2001

- Diagnosis and Management of Dental Caries Throughout Life

2002

- Endoscopic Retrograde Cholangiopancreatography (ERCP) for Diagnosis and Therapy*
- Management of the Clinically Inapparent Adrenal Mass (“Incidentaloma”)*
- Management of Hepatitis C: 2002
- Symptom Management in Cancer: Pain, Depression, and Fatigue*

2003

- Total Knee Replacement

2004

- Improving End-of-Life Care*
- Preventing Violence and Related Health-Risking Social Behaviors in Adolescents*
- Celiac Disease

2005

- Management of Menopause-Related Symptoms*
- Manifestations and Management of Chronic Insomnia in Adults*

2006

- Tobacco Use: Prevention, Cessation and Control*
- Multivitamin / Mineral Supplements and Chronic Disease Prevention*
- Cesarean Delivery on Maternal Request*

2007

- Prevention of Fecal and Urinary Incontinence in Adults*

2008

- Management of Hepatitis B
- Hydroxyurea Treatment for Sickle Cell Disease

2009

- Diagnosis and Management of Ductal Carcinoma In Situ (DCIS)*
- Family History and Improving Health*

2010

- Inhaled Nitric Oxide Therapy for Premature Infants
- Preventing Alzheimer's Disease and Cognitive Decline*
- Vaginal Birth After Cesarean: New Insights
- Lactose Intolerance and Health
- Enhancing Use and Quality of Colorectal Cancer Screening*

2011

- Role of Active Surveillance in the Management of Men With Localized Prostate Cancer*

2012

- No Conferences Held In 2012

2013

- Diagnosing Gestational Diabetes Mellitus