

# Regulating Experimentation in Research and Medical Practice

PAUL ULHAS MACNEILL

## Introduction

The previous chapter discussed *research* involving human beings; in this chapter I consider *experimentation* on human beings. Both activities are expressions of an inclination that is quintessentially human and something we humans derive satisfaction and benefit from. The inclination to explore and experiment is evident from an early age. Without that drive – and discoveries made through exploration, experimentation, and research – life would not be as rich and rewarding as it is. Yet there is an obvious downside. When we experiment on ourselves, on other human beings, on animals, or on the environment, there is a risk of harm. That risk is inherent, by definition, in an activity that is initiated without knowing what the outcome will be. Human experimentation becomes a difficult *moral* issue when the experiment is conducted by one person, who stands to gain from the experiment, on another person, who bears the risks of the experiment.

Although a great deal of experimentation is conducted internationally, there are few reports of adverse events. Yet experimentation on human subjects has caused harm and death. In 1999, Jesse Gelsinger, who was 18 years old, died in an experiment that aimed to correct a genetic deficiency he suffered from. Inquiries into his death, at the University of Pennsylvania Institute of Human Gene Therapy, revealed failures in protective mechanisms that may have avoided his death (Thompson 2000). Ellen Roche died, aged 24, after ingesting a chemical agent that suppressed the action of nerves that are normally active in breathing. She was a healthy technician who had volunteered for a number of experiments at the Johns Hopkins Asthma Research Center where she worked. Inquiries into her death also identified breakdowns in protective mechanisms. Yet, as the Dean of Johns Hopkins University School of Medicine acknowledged openly in responding to Roche's death: "At a certain point some patient is going to die in clinical trials. . . . There is no question about it" (Steinbrook 2002).

How then can we respond to this difficult ethical quandary: that some people stand to gain from conducting experiments that subject other people to real and sometimes unforeseen risks of harm (Capron 2006: 431)?

### *Experimentation and research*

The Cambridge dictionary defines an experiment as a “test done in order to learn something or to discover whether something works or is true.” In this sense, ordinary medical treatment *may* be experimental, and innovative treatments certainly are. The courts, however, have distinguished experimentation from treatment by the extent of risk to the patients (or human subjects of the experiment) and the relative lack of any therapeutic benefit for them (McNeill 1993: 119). It is experimentation in this latter and more confined sense that I am referring to: the kind of experimentation that offers little or no established benefit to the subject and carries with it some risk of harm.

Research, however, is defined by the intention to collect and publish data. Without that intention, the work may be considered innovative treatment but not research (Glatstein 2001). In other words, an experimental treatment would only be regarded as research if the physician had the intention to collect data and publish the results. Yet an intention to publish is only one of the factors (and not the major factor) relating to the risk of harm. From the patient’s perspective, the critical issue relates to *experimentation* rather than *research*. Furthermore, not all research carries any substantial risk, whereas an experiment (defined as above) may well do so, especially when it has a direct bearing on the physical, emotional, mental, or social well-being of a patient. For this reason the focus of this chapter is on *experimentation* rather than *research*.

The question therefore is whether or not it is reasonable to *experiment*, and what circumstances justify *experimentation*. Yet the question most often asked is whether it is reasonable to conduct *research* on human beings. As a result, research has been closely regulated, whereas experimentation (other than experimentation conducted as a part of a formal research program) has been comparatively free of scrutiny and regulation. Consequently, surgeons and other medical practitioners have, until recently, been relatively at liberty to innovate and experiment in the course of their practice.

### *Distinction between experimentation and innovative treatment*

The reasons for this liberty are in part historical, and partly they derive from assumptions made about the intention of professionals in serving their clients or patients. In medicine, “clinical freedom” is a further justification for experimentation. While it is assumed that medical practitioners are motivated by a desire to benefit their patients when they treat them, that assumption may be misplaced when it comes to experimentation, and especially when an experiment has no possible therapeutic benefit. It is accepted that researchers are committed to successful outcomes from their research; and are motivated by the desire for knowledge, and the benefits that may come from participating in research and publishing their findings. This constitutes a bias that may lead to underestimates when assessing risks to the welfare of subjects of experiments. As a consequence, it has been recognized, at least since the 1960s, that the well-being of subjects may be compromised because it is secondary to the researcher’s major preoccupations.

There is a blurred area, however, when it comes to experiments in the course of therapeutic treatment. Doctors who experiment with innovative procedures in the diagnosis, treatment, and management of their patients are assumed to be acting in

the best interest of their patients. The tendency is to regard innovative therapeutic treatment as an extension of therapy, rather than as experimentation, and to leave the decision about whether it is appropriate to experiment or not to the treating doctor's discretion. The assumption is that doctors, in attempting new approaches, are doing the best for their patients and should be free to exercise their clinical judgment. However, a doctor's intention is not sufficient to protect patients who volunteer for medical experiments. When people test their ideas, whether it is in surgical practice or within "non-therapeutic" research programs, they tend to overestimate the benefits and underestimate the risks that may flow from the new procedure. Furthermore, benefits accrue to medical practitioners when they pioneer a new approach in medicine. These factors can, and do, lead to bias in assessing the potential for harm in innovative medical treatment, and put the volunteer, who is suffering from a medical condition, at risk – as the Gelsinger case demonstrated.

The fundamental issue however, and the central *moral* concern, is the same as in any experiment: when a new procedure is tested on people for the first time, there are inherent risks in that the outcome *cannot* be known in advance. In practice, one group of people carries the risks and, until that experimental procedure is better understood and the practical issues resolved, others are the likely beneficiaries. This distinction, between experiments done as a part of medical treatment and experiments in the course of non-beneficial research, has blurred the focus on this central moral dilemma.

## History of Experimentation on Human Beings

It is apparent that experimentation has been a part of medical practice for as far back as there are historical records. Hippocrates took advantage of the exposed cortex of an injured boy to scratch its surface and to observe corresponding movements in the boy's body. In early Egypt and Persia, doctors were permitted to use prisoners for medical experiments. In 1721, condemned prisoners at Newgate Prison in England were offered a pardon if they took part in experimental smallpox vaccinations. Plague experiments were conducted on 900 condemned prisoners in the Philippines in the early 1900s. During World War II, in the Stateville Penitentiary experiments, "volunteer" prisoners were infected with malaria by mosquito bites and anti-malarial drugs were tested to find an effective prophylaxis for American combatants in the Pacific. Even after the war, prisoners in US jails were routinely used to trial new pharmaceutical agents in an ongoing "war on disease" (Pappworth 1967; Annas et al. 1977; McNeill 1993). In the United States in the 1800s, slaves were put into pit ovens so that heat stroke could be studied, and scalding water was poured over them as an experimental "treatment" for typhoid fever. Crawford Long (an American dentist) tested the effectiveness of ether as an anesthetic agent in the amputation of a boy's fingers. In his own words, he "amputated two fingers of a negro boy: the boy was etherized during one amputation, and not during the other; he suffered from one operation, and was insensible during the other" (Wall 2006).

Throughout history, it has been marginalized people in society, including racial minorities, prisoners, and slaves, who have been most experimented on. These experiments attracted little criticism within "educated" society because of prevailing attitudes.

*Large-scale experimentation*

With changes in predominant theories and methods in the late eighteenth and early nineteenth centuries, disease came to be seen (especially in the hospitals of Paris) as the effect of pathological entities within specific organs and tissues of the body (Ackerknecht 1982: 146). Doctors through Europe, impressed with this new scientific approach, deliberately infected healthy human beings with material taken from sick patients to test theories about the transmission of disease. There was no apparent regard for the victims who were harmed by these experiments. There are horrific reports (from Germany, Russia, and Britain) of people having pus, and other body matter from infected people, applied to cuts, injected into them, or placed in body orifices.

For example, in the early 1800s, Dublin physician William Wallace deliberately cut two boys (aged 12 and 15 years respectively) on their thighs and introduced pus or blood from syphilitic patients into the fresh wounds. Within two months, both boys showed the unmistakable symptoms of syphilis (Katz 1972: 286–7). This was not an isolated incident. In the late nineteenth century, German doctor Ernst Bumm introduced a culture of gonococcal material directly into the urethra of two women, which led to gonorrheal infection in both of them. Another German doctor, E. Fraenkel, introduced the secretions of gonorrheal patients into the eyes of newborn babies, who were suffering from other medical conditions from which they were likely to die. One of these babies contracted the disease and died 10 days later, exuding pus from a gonorrheal infection of the eyes. Another German, Dr Tischendorff, performed similar experiments with young children. In 1875, Dr Voss, a Russian physician, injected breast milk from a woman suffering with syphilis into three relatively healthy girls (of ages 13, 15, and 16), two of whom subsequently contracted syphilis. The doctor claimed that the girls were prostitutes and that they had given their consent, yet from any moral perspective, neither their consent nor their social status justified his cruelty (Katz 1972: 285–90).

Fortunately, not all doctors at the time were as callous and uncaring of the plight of the people experimented on. French virologist Viday de Cassi, who himself deliberately infected patients with syphilis (apparently with no qualms), complained that some of his peers refrained from this “greatest service to science” on the ground that they regarded these experiments as immoral (Katz 1972: 289).

A major change in attitude worldwide came as a result of revelations about experiments that had been conducted on human beings in Nazi concentration camps.

*German experimentation in World War II*

In the aftermath of World War II it became apparent that German doctors had conducted experiments with callous disregard for any suffering, harm, or death they inflicted on their human subjects.

The “Allied Control Council” (comprising the United States, Soviet Union, United Kingdom, and France) held trials of Germans accused of war crimes, and empowered the US military to conduct “subsequent Nuremberg Trials.” In the first of these (known as the “Doctors’ Trial”), 20 medical doctors and 3 non-medical personnel were

accused of war crimes and crimes against humanity (as defined in Article 11 of Allied Control Council Law No. 10). Most of the 23 had held positions in various medical services within the Third Reich. Of the defendants, 16 were found guilty, 7 of whom (including Karl Brandt, Hitler's personal physician) were hanged.

The defendants were accused of conducting medical experiments predominantly on Jews, but also on Gypsies, Slavs, the mentally insane, and captured members of the Allied armed forces. They were accused of murder, brutality, cruelty, and atrocities in the course of those experiments. Telford Taylor, the chief prosecutor for the United States, stated in his opening address that the defendants had treated their fellow human beings as "less than beasts" and produced ample evidence to support this charge (Annas and Grodin 1992: 67–93).

Many of the experiments conducted within the Nazi program had a military objective, including "high-altitude" experiments that were designed to test the limits of human endurance and human existence in low-pressure chambers. In other experiments, humans were held in tanks of iced water for up to three hours as a means of testing various methods for resuscitating pilots who had been severely chilled or frozen after falling into the sea. Experiments with typhus, malaria, jaundice, spotted fever, and wounds (deliberately inflicted and infected) were designed to find cures to combat diseases troubling German occupation forces. Experiments in the removal of bone, muscle, and nerves from one group of prisoners, and the transplantation of this material into others, were conducted with the ultimate aim of assisting injured soldiers (Katz 1972: 292–306).

There is an even uglier side. Some of the experiments were part of the Third Reich's program for "racial hygiene" that aimed to "purify" the German people by exterminating and sterilizing unwanted groups. Tests on prisoners were made of various methods for sterilizing men and women. Other prisoners were given poison in their food and observed as they died; or they were murdered, after ingesting poisoned food, and their bodies dissected. Others were shot with poisoned bullets.

The accused argued in their defense that they had acted on superior orders, that the sacrifice of a few lives was necessary to save the lives of many, and that experimentation was necessary to support the war effort. They also argued that much of the experimentation on human subjects throughout history had been conducted in an ethically questionable manner (including experiments conducted in the United States); and that "volunteers" in these medical experiments had seldom given proper consent to take part. While none of these defenses could justify the horrors committed in the name of science, it has to be acknowledged that their claims about the unethical conduct of experimentation on human subjects through history were justifiable.

### *The Nuremberg Code*

Neuro-psychiatrist Leo Alexander, physiologist Andrew Ivy (both from the USA), and German medical historian Werner Leibbrand advised the Nuremberg Tribunal on relevant codes of ethics, including the Oath of Hippocrates, German Codes (of 1900 and 1931), and principles formulated by the American Medical Association for the Nuremberg Doctor's Trial (Annas and Grodin 1992; Shuster 1997). In their judgment,

the American judges enumerated 10 “basic principles” that must be observed in conducting medical experiments on human beings, “in order to satisfy moral, ethical and legal concepts.” These principles adopted much of what had been recommended by Alexander and Ivy but gave added emphasis to an absolute right of a subject to consent to an experiment before it could be conducted. Consent included the right to be fully informed and free of any coercion. The judges went further than the recommendations put to them by adding a right for subjects to bring an experiment to an end (Shuster 1997). Their principles put an onus on experimenters to be scientifically qualified; to justify experiments in terms of potential “fruitful results”; to design their experiments properly and base them on previous animal experiments; to avoid unnecessary physical and mental suffering and injury; and to terminate an experiment if it becomes clear that harm will result. These principles subsequently became known as the “Nuremberg Code.”

In historical terms, this was a formulation that went beyond most previous medical codes in recognizing a difference between patients, within a doctor–patient relationship, and subjects of experimentation, in that the primary goal of an experiment is not treatment but the acquisition of knowledge, regardless of the subject’s best interest. The Nuremberg principles were unique in giving subjects themselves the right actively to protect themselves. All previous codes were based on doctors’ responsibility to protect subjects in their experiments.

### *Japanese medical experiments*

The Germans were not alone in conducting cruel and inhumane experiments on human subjects. From 1932 until the end of World War II, Japanese doctors and biologists conducted horrific experiments, largely on Chinese residents and prisoners and also on some prisoners of war, in a number of heavily guarded installations throughout Manchuria in China. Among other horrors, the experimenters performed live vivisections on men, women, and children with no anesthesia. Some victims had limbs successively frozen and removed until only their heads and torsos remained. Even then they were subjected to experiments with plague and other pathogens. Others were burned, shot with shrapnel, exposed to lethal doses of X-rays, or spun to death in centrifuge machines (Byrd 2005; Harris 2002).

The Japanese experimenters displayed a similar attitude to their victims as the German doctors had to victims in German concentration camps. They were regarded as less than human. For example, staff in Unit 731 referred to the people they experimented on as “Maruta” – a Japanese word meaning “log of wood.”

The Japanese atrocities had little influence on subsequent developments however. This was because they were kept secret for many years after the war to keep from public view an agreement between the American government (at the behest of the American Occupation Forces Command in Japan) and the Japanese experimenters, which gave the experimenters immunity from prosecution if they provided the results of their experiments on human beings. The US Command considered the information was relevant to biological warfare. Most of the information about Japanese experiments only became public as the result of freedom of information actions in the United States in the 1980s (Harris 2002; Williams and Wallace 1989).

*Political dimension*

The stark contrast between the action of the United States against the principal German experimenters and their lack of response to the Japanese counterparts illustrates the important role that politics has played in the development of codes of ethics for the conduct of experimentation on human subjects. Secrecy denied the public an opportunity to react to Japanese atrocities (and US complicity in concealing them) and helped to maintain a belief that atrocities of this sort were an aberration that could be confined to the peculiar circumstances of Nazi Germany. This diverted world attention from the extent of inhumane experimentation worldwide, and the need to give attention to humane standards for experimenting on human subjects in all countries.

## Regulation of Human Experimentation

Although the Nuremberg Court condemned, in the strongest possible way, inhumane experiments on human beings, it was not itself very influential. The attitude of the medical profession was that the circumstances of experimentation in German concentration camps during the war bore little relation to normal medical practice and research in peacetime, and that the Code threatened medical progress (Howard-Jones 1982). The major difficulty the profession had was that an absolute requirement for consent, prior to any experimentation on human subjects, ruled out experiments with children, those mentally incompetent to consent, and unconscious patients. The Nuremberg Code was perceived as a “rigid set of legalistic demands” that challenged the right of doctors to conduct research (Beecher 1970: 279).

Subsequently, in Rome in 1954, the World Medical Association (WMA) adopted “Principles for those in Research and Experimentation” that allowed for proxy consent for experiments on patients who lacked the capacity to consent for themselves. The predominant thrust of the WMA principles was to give primary importance to the responsibility of the researcher rather than the willingness of the subject. It was also to emphasize “therapeutic research” and distinguish that from “non-therapeutic research” (McNeill 1993: 44). A further effect was to give emphasis to *research* rather than *experimentation*.

*Unethical experimentation in the USA*

Ivy, the American College of Surgeons’ expert witness, had testified in the Nuremberg trial that the “principles” he advocated reflected common research practice within the United States (Shuster 1997). Subsequent revelations of American unethical experimentation undermined that assertion.

It was revealed that in 1963 doctors had injected live cancer cells into elderly debilitated patients in the Jewish Chronic Disease Hospital. In 1966, an article by Henry Beecher in the *New England Journal of Medicine* drew attention to 22 unethical experiments that had endangered the lives of human subjects (Beecher 1966). In one of these, intellectually disabled children at the Willowbrook State School were intentionally infected with hepatitis. Also in the late 1960s, publicity was given to the Tuskegee syphilis case

in which 400 poor black men from rural areas in the South, diagnosed with syphilis, were left untreated as a part of a study that began in 1932 to chart the development (“natural history”) of the disease in those men. Even after penicillin, which is an effective treatment for syphilis, became available (in the 1940s), these men were offered no treatment and were simply observed as their condition deteriorated. They had not been informed of their diagnosis (only that they had “bad blood”), nor had they been asked for their consent to take part in the study (Jones 1981).

### *Development of committee review in the USA*

One of the early responses was a requirement in 1966, by the US Surgeon-General, for a committee review of applications for Public Health Service grants. Each applicant was required to state that a committee had considered the risks of the research for any human subjects and had satisfied itself of the adequacy of protection of their rights.

Subsequently, the US Senate established a national commission on human experimentation that published reports, including the Belmont Report (outlining basic ethical principles) and the Institutional Review Board Report with a survey of the Institutional Review Boards (IRBs). The Senate also insisted on the promulgation of regulations covering research on human beings and those regulations incorporated and strengthened the National Institutes of Health policy requiring all publicly funded research to be approved by a committee.

### *Committee review in other countries*

In 1966, Canada followed the United States’ lead and adopted a requirement for review by committee, and the following year the Royal College of Physicians of London recommended committee review of research proposals within its guidelines for research with human subjects. New Zealand introduced a requirement for committee review in 1972 and a policy of committee review was adopted by the National Health and Medical Research Council in Australia in 1973. Other countries also followed this lead, and international codes, such as the Council for International Organizations of Medical Science Guidelines, adopted committee review of research proposals as a major protection for human subjects of experimentation.

Within Scandinavia, Sweden has required ethics committee review from the late 1960s, Denmark since the late 1970s and Finland since the early 1980s. In 1984 the Swiss Academy of Sciences recommended advisory bodies on experimentation and, in the same year, both the Netherlands and Belgian governments issued decrees requiring ethics committee review (McNeill 1993).

## Guidelines, Regulations and Directives to Regulate Human Experimentation

The rules applying to research in most countries have been issued as guidelines (variously called “codes,” “guidelines,” “statements,” or “standards”). These guidelines have been issued by funding bodies such as the Canadian Medical Research Council,



the Australian National Health and Medical Research Council, and governmental departments such as the Departments of Health in New Zealand and the UK and professional bodies such as the Royal College of Physicians of London. Internationally there is very little legislation specifically addressing experimentation on human subjects. The US was exceptional in issuing governmental regulations covering human experimentation. As indicated above, both the Netherlands and Belgian governments have issued decrees. Since that time, the European Parliament has issued directives (Capron 2006: 433).

### *Common features of guidelines to regulate research*

There are a number of common features in all these guidelines, standards, and regulations. They all rely on prior review by a committee of proposals for research on human subjects; the committees are required to consider whether proposals are ethical; they can approve a proposal, request modifications, or reject it; membership of the committees is specified to include some members with expertise in research as well as one or two community (or lay) members. There are some differences between countries about other members of the committees. Typically, the guidelines state the matters that the committee should take into account in deciding whether or not to approve the proposal: for example, the requirement to consider whether the potential benefits of the research justify any risks of harm to the human participants. Although there are few explicit sanctions against researchers, or their institutions, for failure to comply, rejection by an ethics committee, or failure to present a proposal for approval by a committee, will often have implications for funding and may lead to refusal by a journal to publish any results from the research program.

### *Critique of research regulation*

Committee review of research has come in for considerable criticism, particularly from researchers, who argue that it adds considerably to the burden and cost of research administration, slows research down, and deters some research altogether. The claim is that an enormous effort is expended by many people in reviewing research proposals, most of which entails very little risk of harm, and there is little or no gain for all this effort in terms of actually avoiding harm.

Capron's (2006) more telling criticism is that those "elaborate rules and processes" have had the effect of normalizing human experimentation and avoiding the "moral dilemma that lies at the heart of every research encounter": that a person is asked to volunteer for research and accept unforeseen risks of harm in order that others might benefit. He suggests that researchers have a vested interest in enrolling human subjects into research programs and may be unwilling or incapable of a frank conversation that adequately presents this central issue.

### *Therapeutic misconception*

There is a particular problem with clinical research in that desperately ill patients who volunteer may be motivated more by a desire to survive than by altruism (Horrobin

2003). Even when told explicitly that a procedure is experimental and that they may not benefit from it, these patients tend to receive an invitation to enroll in a study as grounds for hope of a cure or remission. This phenomenon has been termed the “therapeutic misconception,” and its ethical significance has continued to be debated in the literature since the term was coined in 1982 (Appelbaum and Lidz 2006; Appelbaum et al. 1982; Miller and Joffe 2006). However, the difficulty is not only on the patients’ side. Medical practitioners find it difficult to adequately inform patients that clinical trials are experiments (Brown, Butow, Butt, et al. 2004; Brown, Butow, Ellis, et al. 2004; Capron 2006). It is claimed that researchers exploit the therapeutic misconception. This is one of the reasons given for arguing that large clinical trials of drugs for people with rapidly advancing diseases are usually unethical (Horrobin 2003).

## Regulation of Experimentation in Surgery and Clinical Medicine

At the outset of this chapter, it was claimed that the important ethical issue relates to *experimentation* on human subjects. However, review by committees is of *research*. The language has shifted to the “ethics of research with human participants” and away from “human experimentation.”

### *Distinction between “therapeutic” and “non-therapeutic research”*

The right to experiment within medicine has been staunchly defended by the medical profession on the ground that “desperate” measures could be taken with seriously ill patients if they offer hope of recovery. The 1954 WMA principles allowed that “operations of a daring nature” could be conducted on sick patients in rare and “desperate cases” if the conscience of the doctor would allow it. In the subsequent WMA Helsinki Declaration of 1964 this was spelt out as a fundamental distinction between “clinical research combined with professional care” and “non-therapeutic clinical research.”

Although this distinction has been challenged, it remains in the current Helsinki Declaration. Doctors are permitted to experiment with new procedures in the care of their patients, especially where “proven prophylactic, diagnostic and therapeutic methods do not exist or have been ineffective.” The Declaration provides that a physician is “free to use unproven or new prophylactic, diagnostic and therapeutic measures, if in the physician’s judgment it offers hope of saving life, re-establishing health or alleviating suffering” (Declaration of Helsinki 2000: Principle 32). This is a continuation of a longstanding practice whereby doctors have offered experimental treatment when it “offers hope” with no requirement that the hope be grounded on anything more substantial than “the physician’s judgment.” This is very different from requirements for the ethical approval of research.

This distinction, between experiments done as a part of medical treatment and those done in the course of non-beneficial research, blurs the focus on a central moral dilemma (as was claimed above). The issue is that patients who are experimented on carry the risks of, and are unlikely to benefit from, new procedures. If the procedure is shown to be effective, it has usually been others who have benefited, when it has become better understood and the practical issues have been resolved (Horrobin

2003). Furthermore, when the assessment of potential benefit is left entirely to the “physician’s judgment,” and the sole criterion is that it “offers hope,” then any procedure, even one that has only a remote possibility of success, can be justified.

Similar latitude to that of the Helsinki Declaration is extended in some national guidelines. For example, Australian ethical guidelines provide that the decision on whether a change in treatment is an innovation or clinical research is “generally a matter for the responsible clinician’s judgment” (Australian National Statement 2007). This sidesteps the important moral consideration of whether or not the proposed change is an *experiment*, and what safeguards are in place to reduce any risks of harm to the patient. It also maintains a culture in which doctors have felt free to experiment on their patients.

### *Nineteenth-century operations on slave women for vesico-vaginal fistula*

There are many historical accounts of experimental surgical procedures that were performed on people suffering from appalling diseases because an experimental procedure offered some hope, however slim. For example, US surgeon Dr J Marion Sims conducted experimental surgery between 1845 and 1849 on seven black slave women, without anesthesia, in an operation to correct vesico-vaginal fistula. This is a painful condition in which a narrow, often ulcerated, channel extends from the bladder to the vagina and allows a continuous discharge of urine. In the course of perfecting an operation to correct this condition, Sims practiced on these women (up to 30 times on one of them) and subsequently offered the operation to white women in the North. He has been accused of performing “unethical experimentation” on “African-American women [who] were enslaved as his experimental subjects” (Ojanuga 1993). A counter-claim is that Sims was motivated by a genuine desire to help black women who had pleaded with him to relieve their symptoms (Wall 2006).

### *Twentieth-century hemodialysis, heart, and transplant surgery*

The twentieth century witnessed many advances in surgery and medicine. Hemodialysis, which now saves the lives of many, was developed in the 1940s. Operations on “blue babies” for heart defects, open-heart surgery, heart and heart-lung transplants, and kidney, liver, and bone marrow transplants all became “success stories” and are now practiced almost routinely. However, many of the people who were first operated on died and, in the course of their treatment, endured enormous suffering.

Thorwald, a German journalist, recounted the experiences of men, women, and children who “submitted to pioneering operations on failing vital organs.” He gave a graphic description of Dutch surgeon Willem J. Kolff’s experiments in the 1940s, in which the blood of patients with diseased kidneys was passed through 20 meters of cellophane tubing in an artificial kidney machine to extract urea. One of the many problems was that the cellophane tore easily, allowing blood to mix with the saline solution in the apparatus. Thorwald described a macabre scene of bloody foam spilling over the side of the apparatus and onto the floor around Kolff and his assistants, who stood in watertight shoes on bricks to avoid getting their feet wet. In the early days of this experiment, 16 patients died, before Kolff’s machine successfully treated a 67-year-old woman suffering with extreme kidney failure (Thorwald 1971: 73–91).

There are similar stories in the development of pioneering operations on congenital heart disease in children, and open-heart surgery. In the 1950s and 1960s, operations performed by Dr Harken to assist faulty heart valves with various artificial devices led to 22 or more deaths, although 10 patients survived with “some improvement.” The first patient to receive a complete mechanical heart valve-replacement died, as did a further 10 patients. However 5 patients did survive this operation. One of the early recipients, Mary Richardson, lived for many years, although she needed two further valve-replacement operations, and suffered a stroke at age 41 that paralyzed the left side of her body. Since those pioneering operations, replacement valve surgery has become a reliable and a relatively safe operation. (Thorwald 1971: 51–70).

Following early lung and heart transplants, most patients survived the initial transplant and lived for a short period, before dying from various complications. In 1963, surgeon James D. Hardy conducted the first lung transplant into a convicted murderer, 58-year-old John Russell, in the Mississippi Medical Center Hospital. Russell survived the operation but died 18 days later from kidney failure. In the following year, Dr Hardy transplanted a chimpanzee’s heart into Boyd Rush, a deaf-mute man, who died on the operating table. In 1967, South African surgeon Christiaan Barnard replaced the diseased heart of 55-year-old grocer, Louis Washkansky, with the healthy heart of a young woman who had been struck by a car and suffered fatal brain injuries. Prior to the operation, Barnard gave Washkansky an unfounded and misleading assurance that he had an “80-per-cent chance.” Washkansky survived in very poor condition and, in spite of persistent efforts by Barnard, including massive doses of anti-rejection drugs, his new heart failed 18 days after the surgery (Thorwald 1971: 217–89).

Barnard was criticized by colleagues for embarking on a human heart transplant even though dogs he had experimented on had only stayed alive for a few days after heart transplants. In justifying himself, he said that the patient with an irreversible disease “will beg you for it. He’ll beg for the chance. Because that’s what it means to him – a chance” (Barnard and Pepper 1970).

### *Criticism of “dying” justification for experimental treatment*

George Annas has maintained a consistent criticism of poorly substantiated surgery and medical treatment that is offered to a dying person who has “nothing to lose” because it exposes terminally ill patients to a “special risk for exploitation.” The patient’s “dying status itself” is used “as an excuse to justify otherwise unjustifiable research” (Annas 1985). In a more recent article (2007), he wrote that, “for seriously ill patients, fear of death will predictably overcome fear of unknown risks.” Yet there is an obligation to protect these patients “because terminally ill patients can be harmed and exploited” and because “there are better and worse ways to die.” Annas had previously described the “horrible and prolonged” deaths of Barney Clark and William Schroder following their artificial heart transplants (1992: 130).

Annas claims that it is an abdication of professional responsibility to rely solely on patient autonomy to justify risky and unsubstantiated experimental treatment. As he puts it, “respecting patient autonomy does *not* require that we accept demands for mistreatment, experimentation, torture, or whatever the dying might want” (1992: 130). On the other hand, surgeons and physicians experience “extreme

difficulty” in responding to desperately ill patients who want any innovative procedure, whatever the risk. Refusal, in the face of repeated demands, is difficult to justify in that the surgeon may be deciding for the patient that it is “better to die” (Little 2008).

What is needed is independent assessment of the proposed innovative treatment. Annas states that “choices can and should be limited to reasonable medical alternatives, which themselves are based on evidence” (2007: 413). However, what counts as a “reasonable medical alternative” and as “evidence” is not easy to specify. Is evidence of successful trials of innovative treatments in animals a sufficient basis for their use in humans? For example, the results of animal trials are taken as evidence of the viability of innovative surgery, and as a basis for approving trials of new drugs with healthy volunteers. Yet animal trials can at best be indicative of the risks of harm and potential for benefit in humans. This is a point that will be expanded in the “Discussion” section below.

Cardiothoracic surgeon Elliot Shinebourne (1984) deemed “haphazard experimentation by many different surgical teams” as unethical. He argued for control over the exercise of clinical freedom by surgeons and proposed that “new operations” should be “subject to the same ethical review as other research procedures.” Plastic surgeon C. M. Ward (1994) criticized the “haphazard and cavalier fashion in which new surgical techniques are allowed to be introduced” and compared the lack of regulation of innovative surgery unfavorably with the “rigorous control demanded of a new drug.”

### *Bristol Royal Infirmary case and Kennedy Report*

Shinebourne’s paper had specifically criticized the “arterial switch procedure in children” in the hands of inexperienced surgeons and the willingness of experienced surgeons to attempt variations. He claimed that these practices led to mortality rates of 52 percent. This was eerily suggestive of the later Bristol Royal Infirmary Case in which 29 out of 53 children (55 percent) who were given arterial switch operations and other procedures died, and another four sustained severe brain damage. Two cardiac surgeons and the Chief Executive of the Royal Infirmary were subsequently found guilty of serious professional misconduct and were struck off the medical register. This case has been described as bringing about a “sea change in medical and wider British societal attitudes to professional self-regulation, clinical competence” (Walshe and Offen 2001). The final report of a public inquiry into this case, conducted by Ian Kennedy, recommended inter alia that:

- any clinician carrying out an established procedure for the first time must be directly supervised by colleagues who have the necessary skill, competence and experience;
- any new invasive clinical procedure undertaken for the first time should be shown to be in the patient’s interests and approved by the local research ethics committee;
- patients are entitled to know the extent to which a procedure is innovative or experimental and to be informed about the experience of the clinician who is to carry out the procedure;
- there should be training of surgeons, particularly in new techniques. (Final Report 2001: Recommendations 99–103)

The UK Department of Health responded comprehensively and positively to the Kennedy Report recommendations (Learning from Bristol 2002). It committed itself to “minimising the number of adverse events occurring . . . when a clinician undertakes a procedure for the first time or when new interventional procedures are introduced” and to establishing a National Institute for Clinical Excellence (NICE), independent of the Department of Health, to give effect to these measures.

### *Registration and investigation of new interventional procedures*

Amongst other activities, NICE established a Safety and Efficacy Register of New Interventional Procedures (SERNIP) and assumed responsibility for registering new interventional procedures according to their safety and efficacy. This program prompted the Royal Australasian College of Surgeons (RACS) to develop the Australian Safety and Efficacy Register for New Interventional Procedures – Surgical (ASERNIP-S) (Boult et al. 2002). There are regular reports of the procedures registered within both SERNIP and ASERNIP-S (ASERNIP-S/RACS 2002). In addition, the Royal College of Surgeons of England has published standards for surgeons that require innovations to be approved by an ethics committee and registered with SERNIP (Good Surgical Practice 2002).

In Australia there have also been measures taken at a governmental level. For example, the New South Wales (state) government Department of Health issued a Model Policy for the Safe Introduction of New Interventional Procedures into Clinical Practice to provide a standard process for assessing new interventional procedures (NSW Health 2003). There are also restrictions on new services that operate through the process of accreditation of hospitals in Australia (Minister for Health for NSW 2004). These measures have resulted in changes in practice in some hospitals (at least). For example, the Royal Prince Alfred Hospital in Sydney now requires that the hospital’s ethics committee approve proposed innovative surgical procedures and new services.

In the United States, however, surgeons continue to test new surgical procedures without prior review or ethics approval and without properly informing their patients (Capron 2006: 441). The American College of Surgeons is said to be considering a registration scheme for new interventional procedures (Campbell 2003).

## Discussion

The chapter has focused on *experimentation* rather than *research* on the ground that the primary moral concern relates to unforeseen risks of harm in an experiment, whether or not that experiment is a part of a research program. At the outset, experimentation was defined as procedures that pose a risk of harm to patients, or to human subjects of research, with a relative lack of any benefit offered to them. Most of the critical attention, however, has been given to *research*, including research where there is little apparent risk of harm (such as qualitative research and studies based on questionnaires). Innovative surgery, and experimental treatment within clinical medicine, have had little scrutiny.

The distinction drawn in the Helsinki Declaration between non-therapeutic research and therapeutic research was one factor in this. Doctors were much freer to experiment

when research was combined with therapy. Another factor was the response of the US Congress to research scandals within the United States and the emphasis that the subsequent regulations gave to research. Another factor has been the robust resistance, evident for example in the World Medical Association, to any restriction of doctors' clinical freedom.

Regulation of human experimentation has developed in response to atrocities and scandals, most notably the Nazi experimentation on humans during World War II and research scandals in the United States following the war. While there was growing criticism of surgeons' freedom to experiment on patients, it took another scandal to bring about change. The Bristol Royal Infirmary case, and the subsequent Kennedy Report, have led to a recognition that new and untested medical interventions need to have been properly assessed, both in terms of their efficacy and their potential for harm, and that doctors need training to perform innovative procedures safely. Whereas scandals in the United States led to changes in the review of research, the United Kingdom, prompted by the Bristol case, has driven changes in practice in (and attitudes toward) experimental surgery and medicine. Australian surgeons, and some health authorities, have rapidly followed this lead. Other countries have yet to adopt comparable measures.

What is at stake, as has been recognized by critics both within and outside the medical profession, is that any experimental procedure should be considered in terms of the "reasonable medical alternatives" and be "based on evidence" (Annas 2007). Yet these may be exceedingly difficult to determine. What evidence is sufficient when a procedure is to be tried for the first time in a human being? Animal studies may provide some indication but can never amount to conclusive evidence. Furthermore, what *kind* of evidence is sufficient? Quantitative studies may give some basis for deciding on the likelihood of a successful outcome, but give no indication of a particular individual's experience, or the extent of possible suffering, in undergoing the procedure.

I suggest that independent assessment of the evidence, taking account of the limits of that evidence and a review of all the circumstances in each particular case, is required to substantiate whether the proposed treatment is a "reasonable medical alternative." Those circumstances should include a report of both the doctor's and the patient's expectations and goals, the nature and extent of the patient's understanding of the experiment and possible consequences (including the risks of harm), and the limits of any evidence suggesting a procedure's efficacy (Kerridge 2008). An independent assessment may also alleviate, although not remove, the difficulty for a surgeon or physician in refusing a patient's request for an unsubstantiated procedure. Yet any innovative procedure, even one that has been thoroughly investigated and appears to offer beneficial outcomes, may carry unforeseen risks of harm.

More is required. The communication between a treating doctor and a desperately ill patient is critical. When the best that can be offered is an experimental procedure, there is a need for frank discussion that goes beyond "the facts." Capron refers to physicians finding a "vocabulary of relationships to fill the gaps, the moral silences . . . a language not solely of duties but of hopes and fears, of uncertainties and magical thinking" (Capron 2006: 435). It takes an ethical commitment, and special skill, to present an option in an open and unbiased manner to a person who is desperate to

find a cure or remission from a serious disease, without leaving her with the illusion that an experimental procedure can be offered as therapy. Both doctor and patient need time to discuss their hopes and fears, and time for silence. It may be that a patient's desperation to try anything masks a deep sadness and resistance to accepting her impending death. The opportunity to talk openly with her treating doctor may allow her to move from this position. Equally, a patient may have reached a calm assessment of her situation and choose, nevertheless, to undergo a procedure that offers no substantial (or substantiated) basis for hope. Many standard and accepted medical procedures began in just this way.

## References

- Ackerknecht, E. H. (1982). *A Short History of Medicine*, Rev. edn. Baltimore, MD: Johns Hopkins University Press.
- Annas, G. J. (1985). Baby Fae: and the "anything goes" school of human experimentation. *Hastings Center Report* 15/1 (February): 15–17.
- Annas, G. J. (1992). The changing landscape of human experimentation: Nuremberg, Helsinki, and beyond. *Health Matrix: The Journal of Law-Medicine* 2/2: 119–40.
- Annas, G. J. (2007). Cancer and the constitution – choice at life's end. *New England Journal of Medicine* 357/4: 408–13.
- Annas, G. J. and Grodin, M. A. (eds.) (1992). *The Nazi Doctors and the Nuremberg Code: Human rights in human experimentation*. New York and Oxford: Oxford University Press.
- Annas, G. J., Glanz, L. H., and Katz, B. F. (1977). *Informed Consent to Human Experimentation: The Subject's Dilemma*. Cambridge, MA: Balinger.
- Appelbaum, P. S. and Lidz, C. W. (2006). Re-evaluating the therapeutic misconception: response to Miller and Joffe. *Kennedy Institute of Ethics Journal* 16/4: 367–73.
- Appelbaum, P. S., Roth, L. H., and Lidz, C. W. (1982). The therapeutic misconception: informed consent in psychiatric research. *International Journal of Law and Psychiatry* 5: 319–29.
- ASERNIP-S/RACS (2002). Available at: [www.surgeons.org/http://www.surgeons.org/Content/NavigationMenu/Research/ASERNIPS/default.htm](http://www.surgeons.org/http://www.surgeons.org/Content/NavigationMenu/Research/ASERNIPS/default.htm) (accessed October 2007).
- Australian National Statement on Ethical Conduct in Human Research (2007). Available at: [www.nhmrc.gov.au/publications/synopses/\\_files/e72.pdf](http://www.nhmrc.gov.au/publications/synopses/_files/e72.pdf) (accessed October 2007).
- Barnard, C. and Pepper, C. B. (1970). *One Life*. Sydney: Australasian Publishing Co.
- Beecher, H. K. (1966). Ethics and clinical research. *New England Journal of Medicine* 274/2: 1354–60.
- Beecher, H. K. (1970). *Research and the Individual*. Boston, MA: Little, Brown and Co.
- Boult, M., Babidge, W., Roder, D., and Maddern, G. (2002). Issues of consent and privacy affecting the functioning of ASERNIP-S. *ANZ Journal of Surgery* 72: 580–82.
- Brown, R. F., Butow, P. N., Butt, D. G., Moore, A. R., and Tattersall, M. H. N. (2004). Developing ethical strategies to assist oncologists in seeking informed consent to cancer clinical trials. *Social Science and Medicine* 58/2: 379–90.
- Brown, R. F., Butow, P. N., Ellis, P., Boyle, F., and Tattersall, M. H. N. (2004). Seeking informed consent to cancer clinical trials: describing current practice. *Social Science and Medicine* 58/12: 2445–57.
- Byrd, G. D. (2005). General Ishii Shiro: his legacy is that of genius and madman. MA thesis, Department of History, East Tennessee State University.
- Campbell, B. (2003). Interventional Procedures Safety and Efficacy. Available at: [www.nice.org.uk/page.aspx?o=38300](http://www.nice.org.uk/page.aspx?o=38300) (accessed October 2007).



- Capron, A. M. (2006). Experimentation with human beings: light or only shadows symposium. In: *A world less silent: celebrating Jay Katz's contributions to law, medicine, and ethics. Yale Journal of Health Policy, Law, and Ethics* 6: 431–64.
- Declaration of Helsinki (2000). As amended in the 52nd WMA General Assembly, Edinburgh, Scotland. Available at: [www.wma.net/e/policy/b3.htm](http://www.wma.net/e/policy/b3.htm) (accessed October 2007).
- Final Report (2001). Learning from Bristol: the report of the public inquiry into children's heart surgery at the Bristol Royal Infirmary 1984–1995. Bristol Royal Infirmary Inquiry (July). Available at: [www.bristol-inquiry.org.uk/](http://www.bristol-inquiry.org.uk/) (accessed October 2007).
- Glatstein, E. (2001). What is research? *International Journal of Radiation Oncology, Biology, Physics* 51/2: 288–90.
- Good Surgical Practice (2002). Royal College of Surgeons of England (September). Available at: [www.rcseng.ac.uk/standards/good.html/?searchterm=Good%20Surgical%20Practice](http://www.rcseng.ac.uk/standards/good.html/?searchterm=Good%20Surgical%20Practice) (accessed October 2007).
- Harris, S. S. (2002). *Factories of Death: Japan's Biological Warfare 1932–45 and the American Cover-up*, rev. edn. New York and London: Routledge.
- Horrobin, D. F. (2003). Are large clinical trials in rapidly lethal diseases usually unethical? *Lancet* 361/9358: 695–97.
- Howard-Jones, N. (1982). Human experimentation in historical and ethical perspectives. *Social Science and Medicine* 16: 1429–48.
- Jones, J. H. (1981). *Bad Blood: The Tuskegee Syphilis Experiment*. New York: Free Press.
- Katz, J. (1972). *Experimentation with Human Beings: The Authority of the Investigator, Subject, Professions, and State in the Human Experimentation Process*. New York: Russell Sage Foundation.
- Kerridge, I. H. (hematologist and Director of Centre for Values, Ethics and the Law in Medicine, University of Sydney) (2008). Personal communication.
- Learning from Bristol (2002). The DH response to the report of the public inquiry into children's heart surgery at the Bristol Royal Infirmary 1984–1995. Available at: [www.dh.gov.uk/prod\\_consum\\_dh/idcplg?IdcService=GET\\_FILE&dID=19383&Rendition=Web](http://www.dh.gov.uk/prod_consum_dh/idcplg?IdcService=GET_FILE&dID=19383&Rendition=Web) (accessed October 2007).
- Little, M. (Emeritus Professor of Surgery, University of Sydney) (2008). Personal communication.
- McNeill, P. M. (1993). *The Ethics and Politics of Human Experimentation*. Cambridge, New York, and Melbourne: Cambridge University Press.
- Miller, F. G. and Joffe, S. (2006). Evaluating the therapeutic misconception. *Kennedy Institute of Ethics Journal* 16: 353–66.
- Minister for Health for NSW (2004). Answer to question from Mr Barry O'Farrell to Minister for Health, 1381–Public Hospital Accreditation; Questions and Answers Paper No. 53. NSW Government Information Bookshop: Hansard and House Papers. Available at: [www.parliament.nsw.gov.au/prod/la/qala.nsf/ad22cc96ba50555dca257051007aa5c8/ca25708400173f67ca25704a0009926b1OpenDocument](http://www.parliament.nsw.gov.au/prod/la/qala.nsf/ad22cc96ba50555dca257051007aa5c8/ca25708400173f67ca25704a0009926b1OpenDocument) (accessed October 2007).
- NSW Health (2003). New South Wales Department of Health: Clinical practice – model policy for safe introduction of new interventional procedures. Available at: [www.health.nsw.gov.au/policies/PD/2005/PD2005\\_333.html](http://www.health.nsw.gov.au/policies/PD/2005/PD2005_333.html) (accessed October 2007).
- Ojanuga, D. (1993). The medical ethics of the “Father of Gynaecology,” Dr J Marion Sims. *Journal of Medical Ethics* 19/1: 28–31.
- Pappworth, M. H. (1967). *Human Guinea Pigs: Experimentation on Man*. Boston, MA: Beacon Press.
- Shinebourne, E. A. (1984). Ethics of innovative cardiac surgery. *British Heart Journal* 52/6 (December): 597–601.
- Shuster, E. (1997). Fifty years later: the significance of the Nuremberg Code. *New England Journal of Medicine* 337/20: 1436–40.
- Steinbrook, R. (2002). Protecting research subjects – the crisis at Johns Hopkins. *New England Journal of Medicine* 346/9: 716–20.

- Thompson, L. (2000). Human gene therapy: harsh lessons, high hopes. *FDA Consumer: The magazine of the US Food and Drug Administration* 34/5 (September–October). Available at: [www.fda.gov/fdac/features/2000/500\\_gene.html](http://www.fda.gov/fdac/features/2000/500_gene.html) (accessed November 2007).
- Thorwald, J. (1971). *The Patients*, trans. R. and C. Winston. New York: Harcourt Brace Jovanovich.
- Wall, L. L. (2006). The medical ethics of Dr J Marion Sims: a fresh look at the historical record. *Journal of Medical Ethics* 32: 346–50.
- Walshe, K. and Offen, N. (2001). A very public failure: lessons for quality improvement in health-care organisations from the Bristol Royal Infirmary. *Quality in Health Care* 10: 250–56.
- Ward, C. M. (1994). Surgical research, experimentation and innovation. *British Journal of Plastic Surgery* 47/2 (March): 90–4.
- Williams, P. and Wallace, D. (1989). *Unit 731: Japan's Secret Biological Warfare in World War II*. New York: Free Press.