

2006

When There Are No Randomized Controlled Trials: A Case History of a Controversial Procedure for Metastatic Legal Doctrines

Jeffrey C. Lerner

Diane C. Robertson

Follow this and additional works at: <http://brooklynworks.brooklaw.edu/jlp>

Recommended Citation

Jeffrey C. Lerner & Diane C. Robertson, *When There Are No Randomized Controlled Trials: A Case History of a Controversial Procedure for Metastatic Legal Doctrines*, 14 J. L. & Pol'y (2006).

Available at: <http://brooklynworks.brooklaw.edu/jlp/vol14/iss2/5>

This Article is brought to you for free and open access by BrooklynWorks. It has been accepted for inclusion in Journal of Law and Policy by an authorized administrator of BrooklynWorks. For more information, please contact matilda.garrido@brooklaw.edu.

**WHEN THERE ARE NO RANDOMIZED
CONTROLLED TRIALS:
A CASE HISTORY OF A CONTROVERSIAL
PROCEDURE FOR METASTATIC BREAST CANCER**

*Jeffrey C. Lerner, Ph.D. & Diane C. Robertson **

INTRODUCTION

Throughout the 1990s, judges were in the position of having to make decisions in cases on high-dose chemotherapy (HDC) with autologous bone marrow or stem cell transplantation (ABMT/SCT) without the benefit of sound medical evidence, and were inadvertently unable to provide a societal check on access to an ineffective, and at times, life-threatening healthcare technology. This unfortunate circumstance was not mitigated by the ability of federal judges to qualify expert witnesses under the principles of the *Daubert* decision, despite the growing use of *Daubert* principles in 1990s.¹ The most credible expert witnesses had been proponents of this ineffective medical procedure.

If the concepts of “evidence” in medical and judicial decision making were congruent, adjudication would be less complex. The case we present on high-dose chemotherapy (HDC) with

* Jeffrey C. Lerner, Ph.D., is the President and CEO of ECRI. Diane C. Robertson, B.A., is the Director of Health Technology Assessment Services at ECRI. ECRI is an independent, private non-profit health services research agency designated as an Evidence-based Practice Center by the U.S. Agency for Health Research and Quality and as a collaborating center of the World Health Organization in technology assessment, healthcare risk management, and patient safety. ECRI and its officers and staff adhere to strict conflict-of-interest rules that prevent owning stock in drug and medical device companies and health plans. ECRI HOME PAGE, <http://www.ecri.org/> (last visited Apr. 24, 2006).

¹ *Daubert v. Merrell Dow Pharm., Inc.* 509 U.S. 579 (1993).

autologous bone marrow or stem cell transplantation (ABMT/SCT) for metastatic breast cancer describes a medical procedure that came before the courts frequently in the 1990s. The controversy over this procedure and the set of circumstances surrounding it created a “perfect storm” that diffused an unproven and even dangerous technology to patients before its effectiveness was appropriately studied. The elements in this perfect storm included physician, hospital, industry, and consumer demands. These demands dovetailed with media demand for dramatic life-and-death stories centered on making villains of managed care organizations (MCOs) and victims of patients denied access to certain procedures.

This case illustrates the incongruence and difficulty that judges confront when asked to consider clinical studies that are presented as medical evidence during a proceeding. Judges have been given the responsibility of determining the credibility of clinical studies submitted as medical evidence and whether to admit them as evidence. Courts that tried cases on HDC with ABMT/SCT were not equipped to determine the credibility of the clinical studies submitted to them as medical evidence. The authors of this article, and the organization of which they are a part, were observers of and then active participants in the controversies surrounding this technology throughout the 1990s.

Though perfect storms that include litigation over healthcare technologies can still form today, better tools are available to discern the credibility of medical evidence. In this article, we begin in Parts I-III by presenting this case history and the controversy surrounding HDC. Finally, in Part IV, we propose a solution to aid courts in the future when called to determine the credibility of clinical trial medical evidence.

I. BACKGROUND

Understanding the issues that surrounded the perfect storm over HDC with ABMT/SCT for breast cancer requires some contextual knowledge of breast cancer and the standard treatments used. Breast cancer is described in stages 0 to IV and a patient’s prognosis is linked in large part to the stage of disease at the time

NEED FOR RANDOMIZED CONTROLLED TRIALS 597

of diagnosis.² At one end of the spectrum are stages 0 and 1, and if detected at these stages, more than 95% of the patients can expect to survive for at least 5 years after diagnosis and treatment.³ Effective treatments are available for this stage of cancer. Initial treatment almost always involves surgery to remove the tumor, followed by irradiation of the affected breast. Additional follow-up (adjuvant) therapy includes chemotherapy and/or hormonal therapy to prevent recurrence.⁴

At the other end of the spectrum is metastatic or advanced (stage IV) breast cancer, for which the outlook is fairly grim and effective treatment is more elusive; metastatic breast cancer is the stage of the disease involved in this case history. The 5-year survival rate is an estimated 26%.⁵ Advanced breast cancer is usually treated with various aggressive chemotherapy regimens—typically a combination of alkylating agents, antibiotics and/or antimetabolites. Alkylating agents have been considered to be particularly useful because dose increases are believed to enhance tumor response rates.⁶ But the higher the dosage regimen, the more severe the treatment-related toxicities, which can in some cases cause death. Major toxicities include liver obstruction, cardiac disorders (arrhythmia, heart failure, inflammation), pneumonitis, central nervous system disorders (seizures, neuropathies, meningitis), gastrointestinal system disorders (nausea, diarrhea, ulceration, hemorrhage), urinary tract system disorders, pneumonia, infection, hypertension, and other serious conditions.⁷

² Lawrence H. Sobin & Irvin D. Fleming, *TNM Classification of Malignant Tumors, Fifth Edition*, 80 CANCER 1803-04 (1997).

³ AMERICAN CANCER SOCIETY, BREAST CANCER FACTS AND FIGURES 2005, at 11 (2005), [http://www.cancer.org/downloads/STT/CAFF2005f4PWSecured .pdf](http://www.cancer.org/downloads/STT/CAFF2005f4PWSecured.pdf).

⁴ NATIONAL GUIDELINE CLEARINGHOUSE, GUIDELINE SYNTHESIS: MANAGEMENT OF EARLY STAGE BREAST CANCER (2001), available at: www.guideline.gov.

⁵ AMERICAN CANCER SOCIETY, *supra* note 3, at 11.

⁶ Edward A. Stadtmauer, *Bone Marrow Transplantation for Breast Cancer*, in Barbara Fowble et al., BREAST CANCER TREATMENT: A COMPREHENSIVE GUIDE TO MANAGEMENT, 489-506 (1991).

⁷ HEALTH TECHNOLOGY ASSESSMENT REPORT, ECRI, *High-dose Chemotherapy with Autologous Bone Marrow Transplantation and/or*

A. The Genesis of HDC with ABMT/SCT for Breast Cancer

In the 1980s, a theory emerged that “more” chemotherapy might lead to “better” response and improved survival, despite the fact that no direct evidence was available that any chemotherapy regimen for advanced breast cancer prolonged survival or improved patients’ quality of life.⁸ Since some chemotherapeutic agents exhibited a steep positive dose-response curve,⁹ many in the oncology community inferred that dose escalation would produce a substantially greater response rate, which in turn would improve survival. However, very high doses also posed higher risks to patients by escalating the side effects and impairing or eradicating the patient’s blood-cell-producing system (which includes the immune system). Nevertheless, if the immune system could be quickly restored after HDC, proponents hypothesized that the procedure might be feasible.¹⁰

Clinical researchers devised various ways to restore this system, which involved harvesting the patient’s stem cells for later reinfusion, or transplantation, after HDC, to reconstitute the system. Stem cells could be harvested from bone marrow in a patient’s hip or from a patient’s circulating blood supply. In addition, certain growth factors, known as colony stimulating factors, could be given to try to increase patients’ production of stem cells before the harvest. The stem cells could then be reinfused in the patient through an intravenous tube after the HDC. If the reconstitution of the blood-cell-producing system failed after HDC, the patient would likely die within weeks or months.

Given that the procedure was very high risk, one might expect that rigorous testing would be conducted to determine whether it really worked better than standard chemotherapy. But such testing

Peripheral Blood Cell Transplantation for the Treatment of Metastatic Breast Cancer, 25, 33 (Feb. 1995) [hereinafter *High-dose Chemotherapy*] (report on file with authors).

⁸ Roger P. A’Hern et al., *Does Chemotherapy Improve Survival in Advanced Breast Cancer? A Statistical Overview*, 57 BRIT. J. CANCER 615-18 (1988).

⁹ *High-dose Chemotherapy*, *supra* note 7, at 15.

¹⁰ *Id.* at 24.

NEED FOR RANDOMIZED CONTROLLED TRIALS 599

is not required under any type of federal regulations when using drugs that have already been approved to treat cancer. Thus, an unproven procedure was available outside of the research setting from its inception and the only major barrier to its diffusion was insurance reimbursement. It was considered by many to be a “cutting edge” procedure and the only hope for a disease that had few good treatment options.

Proving the relative effectiveness of HDC compared to standard-dose therapy would require well-designed randomized clinical trials (RCTs) that compared the two treatments in patients with similar medical characteristics (e.g., age, stage of disease). Results from nonrandomized, uncontrolled studies on HDC were being published, but none of them made the necessary comparisons to assure that the published results were meaningful. Meanwhile, the procedure’s popularity grew from the mid-1980s through the mid-1990s. According to data from the Autologous Blood & Marrow Transplant Registry—North America, from 1989 to 1992, the number of HDC with ABMT/SCT procedures for breast cancer more than tripled. By 1993, more than 2,500 procedures had been performed.¹¹ Trials on HDC increased as well, but not of the kind that would yield a definitive answer: 86 studies on HDC with ABMT/SCT were reported in the Proceedings of the American Society of Clinical Oncology’s (ASCO) annual meetings from 1984 through 1993; none were RCTs.¹² Furthermore, significant decreases were seen in the number of patients participating in trials for advanced breast cancer over this period. Patients did not want to participate in trials because they believed that this new “state of the art procedure” was their best hope for surviving, and the procedure was widely available because many oncologists were willing to perform it. For a patient, participation in a randomized controlled trial meant that they might be assigned to the group in

¹¹ Phillip A. Rowlings et al., *Report from the International Bone Marrow Transplant Registry and the North American Autologous Bone Marrow Transplant Registry*, CLINICAL TRANSPLANT 101, 101-08 (1993).

¹² Rowan T. Chlebowski & L.M. Lillington, *A Decade of Breast Cancer Clinical Investigation: Results as Reported in the Program/Proceedings of the American Society of Clinical Oncology*, 12 J. CLINICAL ONCOLOGY 1789, 1789-95 (1994).

the trial that received standard chemotherapy, rather than HDC. Patients wanted what they perceived (and what their treating physicians claimed) was the best, or only, option, even though no evidence existed to demonstrate that it was the better than standard treatment.¹³

B. The Perfect Storm

The unbridled diffusion of HDC in the 1990s led to a perfect storm. Many circumstances coalesced to create this storm. Enthusiasm for performing this procedure was tied to much more than a belief in its efficacy even though no evidence existed to prove its efficacy.

National Breast Cancer Coalition patient advocate and breast cancer survivor, Musa Mayer, aptly described those times in a recent commentary:¹⁴

It took me some time, and a lot of study, to understand the dynamics of what had actually happened in America with bone marrow transplants in breast cancer. And how wishful thinking on the part of patients and oncologists, public pressure, heart-wrenching media stories of desperately ill young mothers, political and legislative mandates for insurance coverage, personal reputations of researchers, and profit margins of hospitals with transplant beds to fill all managed to widely promote a toxic and expensive treatment before there was sufficient evidence of its safety or efficacy.

. . . [t]he prevailing wisdom of the time was that desperate circumstances called for desperate measures. Many women at the time, including my friends vowed to “go out fighting,” rather than have the longer life and gentler death that might have been theirs with conventional treatment. “If I die,” young women would frequently say, “I want my children to know I did everything I could.” One transplant

¹³ *Id.*

¹⁴ Musa Mayer, *When Clinical Trials Are Compromised: A Perspective from a Patient Advocate*, 2 PLOS MED. 1060, 1060-61 (2005).

NEED FOR RANDOMIZED CONTROLLED TRIALS 601

unit actually used this coercive argument as a marketing ploy. Naively, I believed that doctors could be trusted to rely on good evidence, especially for a treatment as toxic and costly as this. Certainly, they would never allow themselves to be misled by partial evidence or a compelling theory—that more is better, or that dramatic tumor response in uncontrolled Phase II trials actually predicted for clinical benefit.

The procedure was performed by the nation's leading oncologists who held important positions. Clinicians were held in high esteem by patients, colleagues and their professional communities. Whether they were at first leading oncologists who performed the procedure or whether they emerged as leaders in the oncology community as a result of performing HDC is difficult to discern. Physician advocates for HDC hailed from renowned institutions such as Sloan Kettering in New York and Duke University in North Carolina. They published articles in prestigious medical journals such as the *Journal of Clinical Oncology*, the *Journal of the National Cancer Institute*, and the *New England Journal of Medicine*. These same leaders also became involved in the contentious insurance issues related to access to the procedure.

Certain experiences of the authors suggest that some of these oncologists also had vested intellectual and financial interests in the procedure. Vested intellectual interests became apparent when we obtained external peer review on the preliminary draft of our 1995 technology assessment on the procedure. Upon learning the results of ECRI's assessment, a renowned oncologist from a recognized cancer Center of Excellence declared that they could not possibly be valid—not because of the methodology we used, but because the results of our analysis led, in her opinion, to the wrong conclusions.¹⁵

Vested financial interests of some oncologists enhanced diffusion of this technology, but the way this occurred was more complex. In addition to the procedure being performed by the most famous oncology researchers at premier institutions, oncologists in

¹⁵ *High-dose Chemotherapy, supra* note 7.

community hospitals also wished to perform the procedure but did not have the wherewithal to do so. The technical expertise and facilities required to care for these patients led some community hospitals to contract with for-profit centers that offered the procedure—centers in which some famous oncologists had a financial stake. Such centers were purely enterprises created to bring in revenue by performing a high-risk procedure whose efficacy was unknown. It was common for a single HDC with ABMT procedure to cost \$150,000 to \$200,000 in the early to mid-1990s, although procedure advocates worked to lower costs by modifying HDC procedures and creating outpatient facilities to perform it.

Another factor in the evolution of this perfect storm was that some oncology professional societies worked with a for-profit company to lobby state legislators to enact state mandates requiring health insurers to pay for HDC with ABMT/SCT. At times these efforts were successful. These mandates were matched by decisions at the federal level by the U.S. Office of Personnel Management to declare that health insurance plans covering federal employees and their families would be required to cover the procedure.¹⁶

Physician enthusiasm for the procedure was, however, only one element in the perfect storm. Breast cancer became a renowned cause in the 1990s, which marked the advent of an era in which patient organizations such as the Susan G. Koman Foundation became nationally known through events such as the Race for the Cure. The National Breast Cancer Coalition also formed and achieved the signal success of obtaining extraordinary amounts of funding for breast cancer research from the U.S. Department of Defense—a very non-traditional funding source for cancer research.

The high public profile of the disease helped to build an emotional response to the procedure with the help of celebrities and the media. Celebrities affected by breast cancer were willing to

¹⁶ ECRI, *The Impact of Inappropriate Diffusion of Technology*, HEALTH TECHNOLOGY ASSESSMENT NEWS (ECRI, Plymouth Meeting, PA) Nov/Dec 1994, at 1, 6.

NEED FOR RANDOMIZED CONTROLLED TRIALS 603

talk publicly about their experiences and, while most cases of breast cancer occur in older women, a disproportionate number of tragic cases in younger women were covered in the news media. This coverage created greater awareness of breast cancer, but also skewed perceptions about the true prevalence and incidence of the condition.

Another element in this perfect storm was the rapid conversion of the nation's health insurance system to a managed-care model. Managed care organizations (MCOs also known as health maintenance organizations or HMOs) were initially considered a socially progressive attempt to merge insurance payment with better controlled delivery and cost-effectiveness of care. In theory, this merger would slow the spiraling costs of healthcare while increasing the quality of care for patients. But the rapid and ragged implementation of managed care at that time led to a strong backlash from doctors, hospitals, and patients. A perception developed that cost was the sole consideration in MCOs' decision making about what to cover. The locus of coverage decision making, which resided with medical directors of MCOs, presented an unacceptable conflict of interest to the public which challenged the legitimacy of managed care. Many different constituencies and factors play important roles in coverage policy making of managed care plans: the press, lawmakers, consumers, healthcare professionals, healthcare industry representatives, government and state mandates, business and contractual obligations of beneficiaries' insurance contracts, and the scientific evidence on the technology in question.

Strengthening the storm were media demands for dramatic life-and-death stories centered on making villains of MCOs and victims of patients denied access to certain procedures. HDC with ABMT/HDC provided the perfect media opportunity. The media presented an extraordinarily one-sided perspective that contributed heavily to the perception that the procedure must be good because it was expensive and MCOs opposed it. *Time Magazine's* cover on January 22, 1996, featured a physician wearing a surgical mask gagging his mouth, along with the bold headlines "Special Investigation. What Your Doctor Can't Tell You. An in-depth look at managed care and one woman's fight to survive."

The final, though not the least important, element in this perfect storm is the issue of medical uncertainty. That is, statistically based studies show how populations, not individual patients, will fare under a procedure. Thus, even if a procedure is rarely effective, when life is at stake it is common for desperate people to hope that they are the exception to the rule—that they will be the one to benefit. This issue is particularly challenging for the legal profession, which most often deals with compelling individual cases and is not usually in a position to judge whether procedures should be performed by the medical profession.

Business contract arrangements among employers, employees, and HMOs also contributed to the storm and became the basis for disputes entering the legal system over access to the procedure. These disputes are well laid out in a body of work produced by Peter D. Jacobson, J.D., MPH et al.¹⁷ One of the key points in this body of work is that the court's reliance on expert witnesses was problematic because the nation's most credible witnesses were proponents of the technology—highly respected oncologists. Other highly respected oncologists who believed differently from the proponents refused to testify for the defense. Why these credible opposing voices refused to testify is a matter of some speculation, but it was clearly difficult to counter a mounting consensus of support for HDC in the profession. Therefore, the defense in these cases often had to rely on the testimony of insurance company medical directors, who with rare exception were not oncologists and also held little credibility with juries because of public sentiment about MCOs.

In the absence of access to credible expert witnesses, MCOs that were sued sometimes mounted their defense of coverage denials for HDC by citing technology assessment reports (also sometimes called systematic reviews) such as the one ECRI published in 1995. Several reports showed that no evidence was available to prove that the procedure was more effective than—or even as effective as—standard chemotherapy. Some health plans

¹⁷ PETER D. JACOBSON, RICHARD A. RETTIG, & WADE M. AUBRY, *LITIGATING THE SCIENCE OF BREAST CANCER TREATMENT* (forthcoming 2006) (manuscript at 43, on file with authors).

NEED FOR RANDOMIZED CONTROLLED TRIALS 605

also succumbed to the public tidal wave and pressure of expensive litigation and approved coverage for an unproven costly procedure.¹⁸

The technology assessments and methods used to analyze data were drawn into the eye of the storm. From the 1980s through the mid-1990s, as the science of systematic review of medical literature was evolving, the notion of pooling biomedical data for meta-analysis of multiple clinical studies was neither fully developed nor generally accepted. And when such analyses were performed, only data from randomized controlled trials (RCTs) were used because RCTs are generally considered the “gold standard” for evaluating the efficacy of medical procedures. However, for many healthcare technologies, gold standard RCTs often are not conducted—only uncontrolled studies are available. So the question for those engaged in systematic review was how to proceed. In the early 1990s, no group in the field of technology assessment had yet explored pooling data from uncontrolled studies—which were the only studies then available on HDC with ABMT/SCT.

Ultimately in 1995, ECRI published a 328-page systematic review on HDC with ABMT/SCT for the treatment of metastatic breast cancer using the only available data, which were lower quality than gold standard RCT data. The methods of analysis and results of this landmark review are summarized below.

II. ECRI’S CONCERNS GENERATE LANDMARK DATA ANALYSIS

The elements of this perfect storm compelled ECRI to undertake a landmark systematic review of the available data on HDC with ABMT/SCT for metastatic breast cancer. ECRI was motivated by its 30-year mission to improve the safety and cost effectiveness of healthcare and its dedication to providing objective, evidence-based information to the healthcare community for informed decision making. Another motivation was the concern

¹⁸ *New Wave of Diffusion of HDC with ABMT for Breast Cancer: An Answer to the Efficacy Question?*, HEALTH TECHNOLOGY ASSESSMENT NEWS, (ECRA, ECRI, Plymouth Meeting, PA) Mar/Apr. 1996, at 1, 8.

that premature diffusion of this procedure outside of RCTs was misleading and harming patients, and that no one was giving patients complete information. ECRI learned how some patients were being misled after conversations with patients who contacted ECRI for information and from patient advocates, such as Musa Mayer, at the National Breast Cancer Coalition. ECRI was aware that the procedure was being presented to desperately ill patients as their last and only hope without being told of its serious risks. Patients were misled to believe that a high tumor response rate to HDC translated to longer survival. It did not. Some women who survived the procedure were trotted out like poster children to provide testimonials of the procedure's efficacy. Women who were denied the procedure were said to have been denied life-saving treatment. Unfortunately, patients who died from the treatment had no voice, and the oncology community was all too ready to minimize or not acknowledge treatment-related deaths.

Notably, many clinical proponents of HDC claimed that there was a "subset of breast cancer patients" that benefited from HDC, yet no one had ever been able to define that subset. ECRI wanted to find out if that subset really existed.

A. Elements of a High-Quality Systematic Review

Systematic reviews have the ability to yield answers that individual studies cannot because a systematic review examines an entire body of evidence, that is, all the relevant studies. Two key elements of a high-quality systematic review are transparency about the methods and data used to conduct the analysis in the review and the comprehensiveness of the review. Transparency is important for reproducibility of results. Comprehensiveness is important so that all relevant data and information are considered to arrive at the conclusions. Comprehensiveness begins with a thorough and exhaustive search for all the relevant published medical literature. Comprehensiveness also refers to thorough testing of the robustness of all the analyses performed.

For its analysis, ECRI undertook the most comprehensive search for data on HDC with ABMT/SCT that any technology assessment organization had undertaken up to that time. ECRI

NEED FOR RANDOMIZED CONTROLLED TRIALS 607

searched 12 databases to identify clinical studies on the procedure. Those searches identified 1544 articles that were retrieved and reviewed to identify trials with relevant data for possible analysis. ECRI found no controlled trials that made head-to-head comparisons of HDC and standard chemotherapy.¹⁹ At the time, several RCTs were in progress but having trouble accruing patients.

1. Problems with the Quality of HDC Evidence

In reviewing the 1544 articles, ECRI identified many critical weaknesses in the oncology literature.²⁰ These weaknesses made many studies unusable because of insufficient, uninterpretable, or biased data. For example, studies often combined results for patients with different stages of disease and other important variables known to affect treatment and treatment outcomes. Data from such studies were uninterpretable. Treatment-related morbidity and mortality data, which are very important given the toxicity of HDC, were reported inconsistently or not at all. During its examination of the evidence, ECRI found that many studies used a euphemism for describing the deaths that occurred within 30 days of treatment as a result of the treatment. These events were often reported simply as “unevaluable patients.”

Another study quality issue was that many studies enrolled only those patients whose disease was known to respond to chemotherapy. These patients were given a course of “induction therapy”—treatment intended to determine whether or not their tumors were chemo-responsive. While selecting patients whose tumors are most likely to respond is certainly legitimate from a clinical perspective, this practice in a scientific study introduces a serious design flaw. It biases results because only patients with the optimal chance for response were entered into many of the HDC studies. Standard chemotherapy studies did not employ such criteria—those studies enrolled patients with optimal and

¹⁹ *High-dose Chemotherapy*, *supra* note 7, at 34-36.

²⁰ *Id.* at 68-71 (noting “dissimilar HDC regimens and dosages,” differing prior treatment for patients, and patient selection as some of the problems in study design).

suboptimal chances of response. ECRI's analysis showed that this practice in HDC studies was positively correlated with reports of improved survival rates over standard-dose treatment. Thus the benefits of HDC were overstated unless the influence of this patient selection bias difference was accounted for.

HDC studies also adhered to no standard definition or regimen for HDC. HDC regimens in studies had dosages ranging from 2 to 10 times those of standard chemotherapy regimens. The combinations of drugs given varied also. ECRI identified dozens of 2-drug and 3-drug combinations in the literature on HDC. Some regimens included administering biologic growth factors known as colony stimulating factors; others did not. These growth factors were intended to stimulate the patient's production of blood cells to ensure an adequate harvest of stem cells for the transplant procedure that would follow HDC.

Another important weakness of many HDC studies was that patient characteristics were vaguely described. Important characteristics such as those in Table 1, *infra*, which are linked to patient prognosis (and appropriate treatment), were poorly reported or not reported at all in many studies.

Finally, proponents of the technology often supported their assertion of the efficacy of HDC by referencing data from meeting abstracts—short summaries of trials that were never published in full in the peer-reviewed literature. Meeting abstracts were of abysmally poor quality with incomplete and vague information—ECRI had to exclude them from analysis.

B. ECRI's Analytic Methods

Given the absence of RCTs, ECRI analysts identified a statistical method to use on data from uncontrolled studies to make indirect comparisons between HDC and standard-dose therapy. Although indirect comparisons have inherent weaknesses, it was the best available option to try to find an answer about whether HDC was more effective than standard chemotherapy. (Since the time of ECRI's analysis of data from uncontrolled studies, the international health services research community has developed

NEED FOR RANDOMIZED CONTROLLED TRIALS 609

standards for analyzing data from uncontrolled studies.)²¹ Given that none of the ongoing multicenter RCTs were expected to accrue sufficient numbers of patients because patients did not want to enter trials when the procedure was widely available, no one knew whether RCTs would ever be completed.

From the 1544 articles, ECRI identified 40 uncontrolled studies of HDC with ABMT/SCT on a total of 1,017 patients that provided sufficient data to pool together for ECRI's analysis. ECRI also obtained unpublished data from the North American Autologous Bone Marrow Transplant Registry (NAABMTR). This group had created a registry of data on HDC with ABMT/SCT from many centers around the world.

ECRI identified standard chemotherapy RCTs with patient groups that had medical characteristics similar to patients in the uncontrolled HDC studies. The literature searches identified 35 RCTs of standard chemotherapy, which represented data on a total of 4,889 patients that could be pooled to make comparisons to outcomes of HDC patients.

After data were pooled on similar patients from HDC and standard-dose studies, analysts used a statistical technique called meta-regression to see which, if any, patient characteristics led to better outcomes with HDC than with standard chemotherapy. In layman's terms, meta-regression provides a way to explore differences in the characteristics of patients to see how any differences in characteristics might affect treatment outcomes. Characteristics are grouped together and run through a model and then regrouped in different ways and rerun through a model to see what the outcomes are for each group of characteristics. This method would enable analysts to define any subset of patients that might benefit from the procedure. See Table 1.

Table 1. Patient characteristics ECRI analyzed to try to identify patient subset that might benefit from HDC/ABMT

- 1. Treatment regimen given (including induction, purging, CSFs)**
- 2. Age**

²¹ Donna F. Stroup et al., *Meta-analysis of Observational Studies in Epidemiology*, 283 JAMA 2008, 2008-12 (2000).

3. **Estrogen receptor status (positive or negative)**
4. **Response to previous hormonal therapy**
5. **Menopausal status**
6. **Previous chemotherapy for metastatic disease**
7. **Previous adjunctive chemotherapy (for nonmetastatic disease)**
8. **Severity of metastasis:**
 - a. **Lung**
 - b. **Liver**
 - c. **Bone**
 - d. **Skin**
 - e. **Soft tissues**
 - f. **Lymphatics**
 - g. **Contralateral breast**
 - h. **Viscera**
9. **Number of sites of metastasis**

For the regression model, analysts used 61 treatment groups from 35 standard chemotherapy RCTs. These groups of patients had characteristics similar to those in HDC studies. Fourteen regression model groups were analyzed for each of 7 outcome variables. See Table 2 below. The outcomes ECRI considered reflect both the outcomes that ECRI considered important as well as outcomes ECRI believed were misleading, but had been used widely in the oncology community to assess efficacy, such as response rates and response duration. ECRI found that no scientific evidence supported the theory that tumor response rates correlate to improvements in survival. In fact, some evidence suggests that tumor response rates and improved survival are not directly linked. ECRI was also interested in quality of life after HDC because of the high toxicity of the treatment, but no studies evaluated this, so ECRI had no data to address this important outcome.

Table 2. Outcomes ECRI Assessed

1. **Complete and partial (objective) tumor response rate***
2. **Median response duration***
3. **1-year disease-free (progression-free) survival**
4. **2-year disease-free (progression-free) survival**
5. **Median overall survival (duration)**
6. **1-year overall survival****

NEED FOR RANDOMIZED CONTROLLED TRIALS 611

7. 2-year overall survival**

***Complete means that no tumor could be detected upon clinical exam or imaging; partial means that the tumor shrank by 50% or more.**

****The standard length of survival time that is considered indicative of remission is 5 years; however, the longest follow-up that was available in HDC was 1 and 2 years.**

C. Early Death Rates

Patients undergoing HDC run a substantial risk of life-threatening infection due to the total suppression of their immune system and severe toxicity of chemotherapy to major organs—heart, liver, kidneys. ECRI analyzed the early death rates over time. These rates were reported in 31 studies from 1984 to 1994. One might expect that over the many years during which a high-risk procedure is performed, clinical experience with the procedure and accumulating knowledge would yield lower death and complication rates. ECRI's review of data over the 10 years from 1984 to 1994 found no trend toward improvement in death rates from HDC. In fact, a slight trend toward an increase in early deaths was seen in the last two years (1993 through 1994) of studies reporting early deaths.²²

Table 3. Early deaths rates from HDC over a 10-year period

HDC Studies Published in Years:	Number of Studies	% Early Deaths (Mean +- SE)
1984 – 1994	31	10.5 +- 1.4%
1990 – 1994	25	9.7 +- 1.7%

²² *High-dose Chemotherapy*, *supra* note 7, at 25, 33.

1991 – 1994	17	9.4 +- 2.1%
1992 – 1994	14	10.5 +-2.4%
1993 – 1994	7	14.4 +- 3.4%
1994	5	16.6 +-4.3%

1. ECRI's Results

Based on published data and the additional data ECRI obtained from bone marrow transplant registry, ECRI found:

- *No evidence of benefit* for HDC/ABMT/ASCR compared to standard chemotherapy *in any group* when patients were matched for important characteristics.
- Evidence of *harm for all outcome measures* except response rate.
- Substantial evidence for decreased median response duration, median survival time, and one-year overall survival for patients given HDC/ABMT/ASCR.
- Treatment-related death rates were not improving over time.
- Patients receiving optimal standard chemotherapy regimens that were available at the time had better outcomes than patients given HDC/ABMT/ASCR.²³

Based on these findings, ECRI publicly made several recommendations:

- Extremely poor quality of oncologic literature must be improved.
- Editors of oncologic journals should ensure that studies adequately report details of patient characteristics and outcome measures, including deaths.
- “Meeting abstracts” can not be considered legitimate sources of results.
- The public and patients should be informed of the absence of demonstrated benefits.
- Patients considering the treatment should be informed

²³ *Id.* at 1-3.

NEED FOR RANDOMIZED CONTROLLED TRIALS 613

of its potentially significant risks.

- The treatment should be limited to active randomized controlled trials only.

III. HDC EPILOGUE

Soon after ECRI's report was released, the *Journal of Clinical Oncology* published the first RCT to report on HDC with ABMT/SCT compared to standard chemotherapy.²⁴ This was a watershed event because this South African trial from University of Witwatersrand Medical School, Johannesburg reported a significantly higher response rate for HDC than standard chemotherapy. At the May 1999 ASCO meeting, the author, Bezwoda, reported continued positive results of follow-up from this trial while four other RCTs from the U.S. and Europe reported disappointing results from HDC compared to standard chemotherapy.

Bezwoda's results led to an on-site audit of his research data by an international committee which reported egregious and unethical discrepancies in the work, including that no participant in the trial had signed an informed consent.²⁵ His trial publication was ultimately retracted in 2001.²⁶

The first full publication of the largest RCT of HDC with ABMT/SCT for metastatic breast cancer to date came from Edward A. Stadtmauer, et al., at the University of Pennsylvania, and was first published online by the *New England Journal of Medicine* in March 2000 (and in the print journal in April 2000).²⁷

²⁴ W.R. Bezwoda et al., *High-dose Chemotherapy With Hematopoietic Rescue as Primary Treatment for Metastatic Breast Cancer: A Randomized Trial*, 13 J. CLINICAL ONCOLOGY 2483-89 (1995).

²⁵ R.B. Weiss et al., *High-Dose Chemotherapy for High-Risk Primary Breast Cancer: An On-Site Review of the Bezwoda Study*, 335 LANCET 999, 999-1003 (2000).

²⁶ *Retraction*, 19 J. CLINICAL ONCOLOGY 2973 (2001).

²⁷ Edward A. Stadtmauer et al., *Conventional-Dose Chemotherapy Compared with High-Dose Chemotherapy Plus Autologous Hematopoietic Stem-Cell Transplantation for Metastatic Breast Cancer*, 342 NEW ENG. L. REV. 1069, 1069-76 (2000).

Stadtmauer found no significant difference in overall survival rate between HDC and standard dose chemotherapy. Several more RCTs eventually published full results over the few years that came to essentially the same conclusions as ECRI did in its 1995 report.

IV. WHAT WE CAN DO TO PREPARE FOR THE NEXT PERFECT STORM: RECOMMENDATIONS

We make two recommendations for judges. First, that participating in training sessions on the nature and use of scientific methods in medical research is valuable for judges. This is based on our experience in presenting the HDC case study to approximately 80 federal and state judges and engaging with them in dialogue about the value of those sessions. The sessions are not intended to turn judges into statisticians or medical researchers. Rather, the intent was to introduce the concepts of medical research enterprise as it exists today, and as it has existed historically. Revisiting the past to review what has taken place and learn how we can better the system in the future is important. Judges involved in the HDC cases of the 1990s could have asked more penetrating questions if they have been exposed both to the historical record and to the techniques for discerning credible evidence that are available today.

Our second recommendation is to develop a bench book for adjudicating medical technology cases. Such a book could provide informational tools for pretrial, trial, and post-trial procedures that involve evidence from health services research. Because so many dynamic cases are in the courts today, the bench book should be underpinned with an ongoing relational database of cases that use health services research as evidence. For example, the database would have fields that categorize the specific disease or ailment that is the subject of the case, the treatment or technology in question, the procedural posture of the case, the type of health services research entered into evidence or referred to as part of the factual record, and relevant precedents. Other data fields may include, where available, the specific type of defendant (e.g., health plan, third-party administrator, employer), whether

NEED FOR RANDOMIZED CONTROLLED TRIALS 615

alternative dispute resolution was utilized, precedents that were cited, and other information. The database would feed development of the bench book itself and provide fodder for research articles for publication.

One section of the bench book would include a glossary and nomenclature of evidentiary terms derived from health services research, illustrated by the case law and amplified by explanation of the sources and the validity of the scientific foundation of the case law would be cited, as appropriate. For example, mention of a “consensus statement” that appeared in a case would be defined, noting whether it was produced by the National Institutes of Health, a clinical specialty society, or a pharmaceutical company, and what the distinctions among these are.

Another section of the bench book would be a procedural checklist. That is, there would be a description of typical adjudication procedures in which evidence for health services research is proffered. This could be subdivided into sections such as the stages of a trial (e.g., status conference, discovery, motion to dismiss). It would discuss the evidentiary tools applied in each of the professional benchmarks. It would note when health services research has been used and also where it might be used in the future.

A third section would include “frequently asked questions” that presiding judges are likely to encounter. For example, “What factors, other than a review of relevant health services research, might go into an insurer’s coverage decision?” Or, “What is the meaning of FDA ‘approval’ for a pharmaceutical or medical device?”

A fourth section of the bench book would contain resources or examples of instructions which judges could offer to a jury to enable them to weigh the medical evidence presented by expert witnesses. For example, there could be a subsection on “exceptional introduction of evidence” including clinical practice guidelines, technology assessments, and other clinical protocols. There could also be a subsection on application and amici briefs that considers the use of evidence that was not proffered at trial.

CONCLUSION

The authors believe that educational and informational tools can make a significant positive impact on the judicial system, the medical system, and on the ultimate beneficiary of our efforts—the patient.