NOT FOR PUBLICATION WITHOUT THE
APPROVAL OF THE APPELLATE DIVISION

SUPERIOR COURT OF NEW JERSEY
APPELLATE DIVISION
DOCKET NO. A-0387-16T1
A-0978-16T1

BRANDI CARL and JOEL CARL,

Plaintiffs-Appellants,

v.

JOHNSON & JOHNSON, JOHNSON & JOHNSON CONSUMER COMPANIES, INC., IMERYS TALC AMERICA f/k/a LUZENAC AMERICA, INC., and PERSONAL CARE PRODUCTS COUNCIL f/k/a COSMETIC, TOILETRY AND FRAGRANCE ASSOCIATION (CTFA),

Defendants-Respondents.

APPROVED FOR PUBLICATION
August 5, 2020
APPELLATE DIVISION

DIANA BALDERRAMA and GILBERT BALDERRAMA,

Plaintiffs-Appellants,

v.

JOHNSON & JOHNSON, JOHNSON & JOHNSON CONSUMER COMPANIES, INC., IMERYS TALC AMERICA f/k/a LUZENAC AMERICA, INC., and PERSONAL CARE PRODUCTS COUNCIL f/k/a COSMETIC,

TOILETRY AND FRAGRANCE ASSOCIATION (CTFA),

Defendants-Respondents.

Argued October 24, 2019 – Decided August 5, 2020

Before Judges Alvarez, Suter, and DeAlmeida.

On appeal from the Superior Court of New Jersey, Law Division, Atlantic County, Docket Nos. L-6546-14 and L-6540-14.

Richard M. Golomb, argued the cause for appellants (D'Amato Law Firm, Golomb & Honik, PC, and Ted G. Meadows (Beasley Allen Crow Methvin Portis & Miles, PC) of the Alabama bar, admitted pro hac vice, attorneys; Paul R. D'Amato, Richard M. Golomb, Tammi Markowitz, and Ted G. Meadows, on the briefs).

Susan M. Sharko and Kaitlyn E. Stone argued the cause for respondents Johnson & Johnson and Johnson & Johnson Consumer Companies (Faegre Drinker Biddle & Reath LLP, and John H. Beisner, Jessica D. Miller, and Geoffrey M. Wyatt (Skadden, Arps, Slate, Meagher & Flom LLP) of the District of Columbia bar, admitted pro hac vice, attorneys; Susan M. Sharko, John H. Beisner, Jessica D. Miller, and Geoffrey M. Wyatt, on the briefs).

Coughlin Duffy LLP, and Nancy M. Erfle (Gordon Rees Scully Mansukhani, LLP) of the Oregon bar, admitted pro hac vice and Michael R. Klatt and Leslie A. Benitez (Gordon Rees Scully Mansukhani, LLP) of the Texas bar, admitted pro hac vice, attorneys for respondent Imerys Talc America (Lorna A. Dotro,

Mark K. Silver, Nancy M. Erfle, Michael R. Klatt, and Leslie A. Benitez, of counsel and on the briefs).

Jared M. Placitella argued the cause for amicus curiae New Jersey Association for Justice (Cohen, Placitella & Roth, PC, attorneys; Christopher M. Placitella and Jared M. Placitella, of counsel and on the briefs).

The opinion of the court was delivered by

ALVAREZ, P.J.A.D.

These matters, scheduled back-to-back, are now consolidated for decision. Plaintiffs Brandi Carl and Joel Carl, and Diana Balderrama and Gilbert Balderrama, brought suit against defendants Johnson & Johnson, Johnson & Johnson Consumer Companies, Inc., Imerys Talc America, and Personal Care Products Council.[1] The complaints sought damages for personal injury from Brandi Carl and Diana Balderrama's development of ovarian cancer, allegedly from their use of Johnson & Johnson's Baby Powder. Plaintiffs' lawsuits were selected to be the first two to be tried in the "talc-based body powder products" multi-county litigation in Atlantic County. On September 2, 2016, the trial court granted defendants' motion to exclude the opinions of plaintiffs' two principal experts on causation, Daniel Cramer and Graham Colditz. On that basis, the court then granted defendants' motions for

[1] Defendant Personal Care Products Council did not participate in the litigation after the filing of an answer.

summary judgment. The matters were stayed pending the Court's decision in In re: Accutane, 234 N.J. 340 (2018).[2] We now reverse.

The trial judge barred plaintiffs' expert opinions after an N.J.R.E. 104 hearing conducted pursuant to Kemp ex. rel. Wright v. State, 174 N.J. 412, 427 (2002). He considered testimony from all the experts, including defendants', as well as extensive submissions by the parties. The judge found fault with "the narrowness and shallowness of [plaintiffs' experts'] scientific inquiries and the evidence upon which they rely. Their peers in the scientific community would not rely upon such limited information." He further found that "their areas of scientific inquiry, reasoning, and methodology, are slanted away from objective science and towards advocacy." He did not believe that their opinions relied upon "'data or information used[] soundly and reliably generated and one of a type reasonably relied upon by comparable experts,'" paraphrasing the language of Rubanick v. Witco Chemicals Corp., 125 N.J. 421, 449 (1991). The judge relied upon his own reading of the supporting material to dismiss the opinions of plaintiffs' principal experts as flawed. In other words, his conclusions went to the merits of their opinions and his

[2] Plaintiffs seek a remand to have the opportunity to present their evidence in terms of Accutane and Daubert v. Merrell Dow Pharmaceuticals, Inc., 509 U.S. 579 (1993), and present newly available scientific evidence. We do not agree such a remand is necessary in light of our decision that the judge incorrectly concluded plaintiffs' experts' methodology was improper.

disagreement with them, rather than their methodology and the soundness of their data. In some instances, he relied upon defendants' expert opinions to explain his disagreement, and mischaracterized it as proof of unsound methods. Since the judge found the experts' methodology suspect, and considered them biased, he suppressed their opinions and granted defendants summary judgment. The judge did not criticize any particular study in the hearing record, including those on which plaintiffs' experts relied, as flawed or otherwise unworthy of reliance.

I.

In Accutane, which all agree applies to this appeal, the Court closely analyzed N.J.R.E. 702 and 703, and our state's application of Daubert v. Merrell Dow Pharms., Inc., 509 U.S. 579 (1993). The Court reiterated that the trial judge's function is to act as a gatekeeper, not to substitute his or her judgment for that of "the relevant scientific community." Accutane, 234 N.J. at 390 (citing Landrigan v. Celotex Corp., 127 N.J. 404, 414 (1992)). The inquiry is whether the experts adhered to "the same level of intellectual rigor that characterizes" their field. Id. at 386 (quoting Kumho Tire Co. v. Carmichael, 526 U.S. 137, 152 (1999)). A trial judge must "focus on the expert's principles and methodology—not on the conclusions they generate." Id. at 384. The critical determination is "'whether comparable experts accept

the soundness of the methodology, including the reasonableness of relying on [the] type of underlying data and information.'" Id. at 390 (quoting Rubanick, 125 N.J. at 451). When a trial court in a civil matter excludes an expert opinion on "unreliability grounds" after conducting "a full Rule 104 hearing," a reviewing court "must apply an abuse of discretion standard" to that determination. Id. at 391.

The judge granted defendants' summary judgment applications dismissing the complaints, after suppressing plaintiffs' expert opinions. A grant of summary judgment is reviewed de novo. Cypress Point Condo. Ass'n v. Adria Towers, LLC, 226 N.J. 403, 415 (2016). We "review the competent evidential materials submitted by the parties to identify whether there are genuine issues of material fact and, if not, whether the moving party is entitled to summary judgment as a matter of law." Bhagat v. Bhagat, 217 N.J. 22, 38 (2014) (citing Brill v. Guardian Life Ins. Co. of Am., 142 N.J. 520, 540 (1995); R. 4:46-2(c)).

We conclude, contrary to the trial judge, that the experts' opinions were indeed based on sound methodology applied to data upon which experts in their field may reasonably rely. Therefore, genuine issues of material fact preclude the grant of summary judgment to defendants. We combine our discussion of the issues raised by plaintiffs on appeal.

II.

We begin, as the Court directed in Accutane, with the analytical structure taken from the Federal Judicial Center's Reference Manual on Scientific Evidence (Third Ed. 2011) (the Manual). Epidemiology and epidemiological studies of various types are "used to test whether exposure to a particular agent causes a harmful effect or disease." Accutane, 234 N.J. at 352-53. The Court explained:

> [T]hree basic questions arise in the assessment of a study's methodological soundness:
>
> > 1. Do the results of an epidemiologic study or studies reveal an association between an agent and disease?
> >
> > 2. Could this association have resulted from limitations of the study (bias, confounding, or sampling error), and, if so, from which?
> >
> > 3. Based on the analysis of limitations in Item 2, above, and on other evidence, how plausible is a causal interpretation of the association?
>
> [Id. at 354 (citing to the Manual at 554).]

"Once an association has been found between exposure to a particular agent and development of a specific disease, researchers then consider whether that 'reflects a true cause-effect relationship.'" Id. at 354 (citing to the Manual

at 597). In making that assessment, certain factors, known as the Hill criteria or Hill factors, guide the determination. Ibid.

Furthermore, the Court clarified that New Jersey courts shall rely upon the Daubert factors when considering the reliability of the scientific methodology. Id. at 398-99. Those factors, "pertinent for consideration, but not dispositive or exhaustive," are:

> 1) Whether the scientific theory can be, or at any time has been, tested;
>
> 2) Whether the scientific theory has been subjected to peer review and publication, noting that publication is one form of peer review but is not a "sine qua non";
>
> 3) Whether there is any known or potential rate of error and whether there exist any standards for maintaining or controlling the technique's operation; and
>
> 4) Whether there does exist a general acceptance in the scientific community about the scientific theory.
>
> [Id. at 398. Cf. Daubert, 509 U.S. at 593-94 (same list of four factors, by which U.S. Supreme Court did "not presume to set out a definitive checklist or test").]

An expert opinion is unreliable unless its proponent can "demonstrate the soundness of a methodology, both in terms of its approach to reasoning and to its use of data, from the perspective of others within the relevant scientific community." Id. at 400.

The Court cited to In re: Rezulin Products Liability Litigation, 369 F. Supp. 2d 398, 425 (S.D.N.Y. 2005), for its admonition against expert reliance on just selective portions of the body of relevant scientific information. Accutane, 234 N.J. at 400. Rezulin held that Daubert requires experts at least to consider contrary evidence. Rezulin, 369 F. Supp. 2d at 425. They must address "obvious alternative explanations" by explaining "information that otherwise would tend to cast doubt on" their theories, because an opinion that "does not acknowledge or account for" such evidence is unreliable. Ibid. The amount of evidence tending to contradict the expert's theory or conclusions may be large enough that ignoring it amounts to selectivity as opposed to adherence to the field's intellectual standards. Id. at 425-26 (citing Kumho, 526 U.S. at 152). In sum, the question to be answered is "whether the scientific community would accept the methodology employed by plaintiffs' experts and would use the underlying facts and data as did plaintiffs' experts . . . ." Accutane, 234 N.J. at 400.

III.

We summarize the principles governing epidemiological studies and their use, and the studies in the hearing record.

The two main kinds of epidemiological studies are cohort studies and case-control studies. Manual at 556. A prospective cohort study enrolls a

study population of exposed and unexposed persons and follows it into the future, while a retrospective cohort study "constructs" a study population as of some prior date and follows it "over historical time toward the present." Id. at 557. A prospective cohort study can have the advantage of being better able to establish "the temporal relationship between exposure and disease." Id. at 558.

A case-control study starts with a set of "cases" who have been diagnosed with the disease, assembles a control group of persons without that diagnosis, and compares them in light of prior exposure to the agent. Id. at 559. Case-control studies "are particularly useful in the study of rare diseases," because a cohort study would require "an extremely large group" in order to contain "a sufficient number of cases for analysis." Id. at 560.

When multiple epidemiological studies have reached different results about the existence of an association or its magnitude, a pooled analysis or a meta-analysis may be performed to determine whether their data would yield meaningful results if analyzed together. Id. at 606-07. Care is needed to account for heterogeneity—the extent to which differences in study design contribute to a greater degree of variance among the individual studies' results than would be expected from chance alone. Id. at 607-08.

The starting proposition of any epidemiological study is that the association of the agent with the effect in question has occurred by chance,

without an actual causal relationship. That is called "the null hypothesis." Id. at 241, 574. The study then proceeds to determine relative risks or odds ratios, whether they are statistically significant, and the likelihood that the associations arose by chance if the null hypothesis is true. Id. at 241. A lower likelihood means a stronger inference that the null hypothesis is not true. Ibid.

As the Manual repeatedly emphasizes, epidemiological studies are statistical exercises, and no set of statistical results is capable of establishing that the null hypothesis is actually true or false. "Probabilities govern the samples, not the models and hypotheses. The significance level tells us what is likely to happen when the null hypothesis is correct; it does not tell us the probability that the hypothesis is true." Id. at 252.

The calculated association typically is expressed as a relative risk ratio in cohort studies and as an odds ratio in case-control studies. Id. at 566-69. They are substantially equivalent for most purposes. Id. at 625; see also id. at 569 n.61. They are often simply called "the association" between the agent and the effect. The subtle mathematical differences between them are not germane here, and none of the experts objected to direct comparisons of relative risk ratios and odds ratios.

Certain conventions are followed in evaluating the strength of the inference about causation that a study's results can support. A relative risk or

odds ratio of 1.0 means that the association is just as likely to arise from chance regardless of whether the null hypothesis is false or true. Id. at 567-69. In other words, it establishes the absence of an association in that study. Ibid. A ratio greater than 1.0 means that an association exists. Ibid.

Another convention is that the study results, whatever they are, must be "statistically significant." Id. at 573. The typical standard is to calculate for statistical significance at the 95% level, id. at 245, 251, which all of the studies and expert analysis in this case applied. Even when the association is greater than 1.0, it is not statistically significant unless the entire range of the 95% "confidence interval" for the association, the range of results that would contain the true association for the study population 95% of the time if the study were repeated, is greater than 1.0. Id. at 247, 579-81. In addition, the value of *p*, the probability that the data showing a relevant match within the population occurred by chance rather than from an actual association, must be sufficiently "small," although the Manual cautions that *p* tends to decrease as sample size increases regardless of whether the actual association is "legally or practically important." Id. at 250-53.

All the experts in this case agreed on those conventions, and on the need for a statistically significant association greater than 1.0 before proceeding to consider the possibility that the association may justify an inference of

causation. They also accepted the calculations in the studies submitted and their authors' representations about the existence or lack of statistical significance. However, for particular studies, the experts sometimes disagreed on whether the relative risk or odds ratio needed adjustment to mitigate a weakness in study design, whether the ratio in a particular study was "strong" or "weak," and more generally, on how far above 1.0 the association needed to be in order to support the author's inferences.

The strength of the inference that can be drawn from an epidemiological study's results is not to be confused with the study's "power." Power is the likelihood that the study will conclude that the null hypothesis is false when it actually is false. Id. at 254 n.106, 582. In more practical terms, power is "the chance that a statistical test will declare an effect when there is an effect to be declared." Id. at 254. Power reflects both the size of the effect and the size of the sample. "Discerning subtle differences requires large samples," while "small samples may fail to detect substantial differences." Ibid.

However, the Manual gives no indication of when a sample size may be considered "small," let alone too small for any particular purpose. It is "[c]ommon sense" that the study population needs to be "large enough," and that enlarging it would allow "a more accurate conclusion and reduce the chance of random error." Id. at 576. Yet "[t]here is no easy answer" to the

question of how large the sample size "should" be, because increasing it would not reduce bias, which is a function of study design. Id. at 246. Furthermore, "beyond some point, large samples are harder to manage and more vulnerable to nonsampling error." Ibid.

Accordingly, in evaluating bias, a study's design must be considered, not just its size. Id. at 583. Selection bias, recall and other information bias, and classification bias can exist in both case-control and cohort studies. See id. at 584-90. "Most epidemiologic studies have some degree of bias that may affect the outcome." Id. at 583. While the bias "can be difficult, if not impossible," to identify, ibid., the strength or consistency of the association "may suggest that a bias, if present, had only limited effect." Id. at 585.

Similarly, both cohort and case-control studies can have confounders, which are events or traits that may cause or contribute to the effect in question independently of the agent being investigated, or conversely, in some correlation with the agent. Id. at 590-91. The influence of confounders can be mitigated, or at least estimated, by a statistical sensitivity or multivariate analysis of the study data and results. Id. at 591-97. One such technique is stratification, the creation of subgroups by specified criteria such as age or extent of exposure to the confounder. Id. at 596-97, 628.

Both sides agreed that, in evaluating an epidemiological study and its results, a statistically significant association even after adjustment for bias and confounders is just the starting point. Accurate rejection of the null hypothesis does not automatically establish any particular alternative hypothesis. Id. at 257. The experts here, and the court, relied on the seminal and still highly influential factors that Sir Austin Bradford Hill proffered on just how an epidemiological study should be evaluated before its reported statistically significant association between exposure to an agent and a disease may be considered support for an inference of a causal relationship.[3]

Hill observed that, for purposes of preventive medicine, "the decisive question" is whether a change in an environmental factor will alter the frequency with which the undesirable event in question occurs. Hill at 295. In other words, a causal relationship must exist, but the extent to which the relationship's mechanism should also be demonstrated before recommending action "will depend upon circumstances." Ibid.

---

[3] These factors appear in the transcription of Hill's address to the Royal Society of Medicine's Section on Occupational Medicine. Austin Bradford Hill, The Environment and Disease: Association or Causation? President's Address, 58 Proceedings of the Royal Society of Medicine 295 (1965), https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1898525/pdf/procrsmed00196-0010.pdf. It is cited here simply as Hill. The Manual recognized Hill's factors and proceeded to a substantially similar discussion of how to evaluate an epidemiological study as support for inferring causation. Manual at 598-603. However, the experts and the court cited only to Hill.

Hill named nine factors to consider in evaluating an epidemiological study for whether it supports an inference of causation. Id. at 295-99. He emphasized that not all of them are required in every instance, and that no single factor is mandatory in all instances. Id. at 299. The first factor was the strength of the association, which needed to be considered in light of all the possible causes of the undesirable event. Id. at 295-96. A strong association may be an appropriate threshold when confounders readily come to mind, but Hill cautioned that confounders should be "easily detectable" before they are used to preclude an inference of causation about the agent in question. Id. at 296. Indeed, he admonished that "[w]e must not be too ready to dismiss a cause-and-effect hypothesis merely on the grounds that the observed association appears to be slight," especially when the event is relatively rare. Ibid.

Hill's second factor was consistency of results, with similar results that were "reached in quite different ways, e.g. prospectively and retrospectively," being the most notable. Id. at 296-97. The third was specificity, which again can be impressive, but cannot be mandated, because "diseases may have more than one cause," or because an agent might be a cause of several diseases. Id. at 297. The fourth, a temporal relationship of exposure to the agent preceding the disease, may pose a question for "diseases of slow development" that might

somehow cause the behavior or exposure that was initially suspected of causing the disease. Id. at 297-98. The fifth, a biological gradient, also called a dose response, can be weighty, although it can only be assessed when it is possible to "secure some satisfactory quantitative measure of" the relevant exposure. Id. at 298.

Hill called his sixth factor, biological plausibility, "a feature I am convinced we cannot demand" because it "depends upon the biological knowledge of the day." Ibid. "[T]he association we observe may be one new to science or medicine and we must not dismiss it too light-heartedly as just too odd." Ibid. However, Hill's seventh factor, coherence, serves in effect as a back-stop on not demanding a biologically plausible mechanism of causation, because it holds that "the cause-and-effect interpretation of our data should not seriously conflict with the generally known facts of the natural history and biology of the disease." Ibid.

Hill believed that his eighth factor, a demonstrated beneficial effect from taking preventive action against the agent in question, might give the most support for an inference of causation, although he noted that such evidence was only "occasionally" available. Id. at 298-99. His ninth and final factor, analogy to the known causal relationship between another agent and disease, would sometimes justify taking preventive action on "slighter but similar

evidence" that the agent in question is analogous in kind and that the disease in question is analogous in severity. Id. at 299.

Hill urged his audience, officials responsible for public and occupational health, to take or decline preventive action only after considering the harm to be avoided, and also considering the possible "injustice" of the costs or intrusions that would be imposed from prohibiting exposure to an agent that did not in fact cause the disease. Id. at 300. The evidence needed to justify such action could be "relatively slight" or "very strong." Ibid. However, he ended with an admonition never to require absolute certainty before acting:

> All scientific work is incomplete - whether it be observational or experimental. All scientific work is liable to be upset or modified by advancing knowledge. That does not confer upon us a freedom to ignore the knowledge we already have, or to postpone the action that it appears to demand at a given time.
>
> [Ibid.]

All the above addresses general causation. Plaintiffs and their experts accepted that epidemiological studies cannot serve as the sole evidence of "specific causation," the proof that a particular plaintiff's disease developed because of the nature and extent of her exposure to the agent in question. However, the Manual, at 608-18, recognizes that epidemiological studies that support general causation may serve to support a plaintiff's burden of

proffering sufficient evidence of specific causation to reach a jury, if due regard is given to the plaintiff's degree of similarity to the study populations in exposure, development of the disease, and other relevant factors.

In 2010, the World Health Organization's International Agency for Research on Cancer (IARC) published volume 93 of IARC Monographs on the Evaluation of Carcinogenic Risks to Humans, which addressed carbon black, titanium dioxide, and talc. It concluded that there was "limited evidence" that perineal[4] talc use could cause ovarian cancer. It noted that "many" case-control studies found a "modest, but unusually consistent, excess in risk," although evidence for dose response was inconsistent, the "impact of bias and potential confounding could not be ruled out," and "the one cohort study" did not support an association. Other reservations were the variety in the studies' definitions of exposure, and the possibility that some of the talc may have contained independently carcinogenic material, like asbestos.

On April 1, 2014, the Food and Drug Administration (FDA) issued a letter in which it denied two petitions to require a warning on consumer talc products that frequent perineal use increases the risk of ovarian cancer. The petitions asserted that talc may contain asbestos, that talc is itself a carcinogen,

[4] The expert witnesses treated perineal use and genital use interchangeably. Any unspecified reference to talc use in the record, including the documentary evidence, refers to such use.

and that epidemiological studies established a causal relationship between genital talc use and ovarian cancer.

The letter stated that the FDA had the authority to propose a regulation with such a warning if a petition for it "is supported by [an] adequate scientific basis on reasonable grounds." However, after reviewing the petitions, responsive comments, and "additional scientific information," the FDA found an absence of evidence that currently marketed talc products might contain asbestos, and a paucity of evidence that talc itself is carcinogenic. The FDA further found that the epidemiological studies the petitioners cited were inconsistent with the ones it located in its own literature searches. It also found study design flaws, which were the failure to confirm that the talc was free of asbestos, and the failure of any one study to address all known confounders including selection and other biases.

The FDA further noted the absence of a "cogent biological mechanism by which talc might lead to ovarian cancer," in light of cases of ovarian cancer that occurred even with no talc exposure, and the lack of evidence for the "incessant ovulation" and "gonadatropin" hypotheses. It acknowledged that the potential of particles like talc "to migrate from the perineum and vagina to the peritoneal cavity is indisputable," which made it "plausible" that perineal talc could migrate to the ovaries and "elicit a foreign body type reaction and

inflammatory responses that . . . may progress to epithelial cancers." The "best evidence for an association or causal relationship" was the epidemiological studies reporting such results, and "the growing body of evidence to support a possible association between genital talc exposure and serous ovarian cancer is difficult to dismiss." Nonetheless, the absence of "conclusive evidence of a causal association" between perineal talc use and ovarian cancer meant that the evidence was insufficient for the FDA "to require as definitive a warning as you are seeking."

As of August 8, 2016, the version for healthcare professionals of the "PDQ" summary titled "Ovarian, Fallopian Tube, and Primary Peritoneal Cancer Prevention," at the website of the National Cancer Institute (NCI) provided an overview of those cancers and possible risk factors. It cited many studies, including some of those in the next part of this opinion. It stated estimates for 2016 of 22,280 new diagnosed cases of ovarian cancer and 14,240 deaths from the disease. As of 2012, the "population lifetime risks" were 1.3% for developing the disease and 0.97% for dying from it. Both figures reflected small but statistically significant decreases during the preceding ten to twenty-five years.

The NCI website focused on epithelial ovarian cancer because it is the most common type. Epithelial cancer comprises the histological subtypes of

serous, mucinous, endometrioid, and clear cell. Those subtypes are "heterogeneous," which suggested that they might arise by "different molecular pathways." Overall, ovarian cancer "is a rare cancer," so if the association of a risk factor with a particular subtype is "moderate," the ability of epidemiological studies to detect it may be "limited" due to sample size and statistical power.

The website characterized risk factors for ovarian cancer as having "adequate evidence" or "inadequate evidence" of an association with an increased or decreased risk of the disease. The evidence was adequate for an increase in risk from obesity and for hormone or hormone replacement therapy, and for a decrease in risk from oral contraceptives, injectable contraceptives, tubal ligation, and breast-feeding. Inconsistent study results meant that evidence was inadequate for a decrease in risk from aspirin and nonsteroidal anti-inflammatory drugs (NSAIDs), as well as for an increase in risk from smoking or perineal talc exposure.

On January 4, 2019, after these appeals were filed, that section of the NCI website was updated. https://www.cancer.gov/types/ovarian/hp/ovarian-prevention-pdq. Although there were several minor changes, the conclusions and the characterizations of the state of the evidence remained the same. The only change germane to this case was the discussion of a May 2016 case-

control study, which was in the record below but not cited by any expert or the court. Ibid. That study, by Joellen Schildkraut and others, Association between Baby Powder Use and Ovarian Cancer in African Americans, 25:10 Cancer Epidem., Biomarkers & Prevention 1411, 1414-15 (2016), comprised 584 cases and 745 controls and found a statistically significant association between genital powder use and the risk of epithelial ovarian cancer. It also found a dose response when study subjects who had ever used talc genitally were compared to subjects who never used it in any manner ("ever user" or "ever use" versus "never user" or "never use"), as well as for daily genital use versus less frequent use. Ibid. The authors considered the results consistent with the causation theory of talc-induced "localized chronic inflammation in the ovary." Id. at 1416. Notwithstanding the addition of that study, all of that section of the NCI website's conclusions and characterizations of the state of the evidence remained the same.

The judge asked the parties to submit the relevant scientific studies and articles cited in their experts' reports or that their experts' testimony would reference. All of the cited studies and articles were published, and neither the court nor any of the experts questioned the merits of their pre-publication selection or review. The relevant ones are summarized here. The court did not criticize any of the studies for having an unsound methodology, for misstating

results, or for failing to consider bias and confounding influences, and it accordingly did not find that the relevant scientific community would consider them unsusceptible of appropriate reliance.

In 1982, Cramer and coauthors published Ovarian Cancer and Talc, 50:2 Cancer 372 (1982), which purported to be the first epidemiological study of talc and ovarian cancer.[5] It was a hospital-based case-control study. The controls were matched to the cases by residence, race, and age. The controls also had to confirm that they still had at least one ovary. The only classification of talc use was "regular," with no indication of duration or frequency.

For cases who used talc on both the perineum and on sanitary napkins compared to never users, the relative risk was 3.28, and it was statistically significant. For all cases, meaning those who used talc in both of those ways or just one, the relative risk compared to never users was 1.92 and still statistically significant. For cases who used talc in only one of those ways, the relative risk of 1.55 was of "borderline" statistical significance. Menstrual history was too homogenous to be a confounder, and adjustments for

[5] Most subsequent references herein to a particular study will be by the lead author's name and the date, for example, "Cramer's 1982 study," or "Cramer 1982" in a parenthetical.

hysterectomy, tubal ligation, parity,[6] and oral contraceptive use did not change the significance of the results.

The authors stated that the link of talc to ovarian cancer was predicated on an analogy to the role of asbestos in mesothelioma, and thus required talc to be able to migrate to "the pelvic cavity," which had been implied by findings of talc particles "embedded in normal and abnormal ovaries." They hypothesized that talc on the ovarian surface could enter an ovary during the foreign body entrapment of ovarian surface epithelium in the inclusion cysts that can form after ovulation, which is the eruption of an ovum through its follicle for travel via the fallopian tube from inside the ovary to the uterus. Alternatively, talc on the ovarian surface might stimulate the entrapment of surface epithelium even between ovulations. The authors concluded that, due to "the histologic and clinical diversity of ovarian cancer, talc exposure is unlikely to be the only cause," and the interaction of perineal talc exposure with other aspects of reproductive tract function merited further study.

In 1989, Bernard Harlow and Noel Weiss published A Case-Control Study of Borderline Ovarian Tumors, The Influence of Perineal Exposure to Talc, 130:2 Am. J. Epidem. 390 (1989). It was a population-based case-

[6] "Parity" means having had a viable pregnancy, even if it did not result in a live birth. Not having had such pregnancies is called null parity or nulliparity.

control study prompted by the "marked differences" in age and survival rates between patients whose epithelial ovarian tumors were borderline and those whose tumors were malignant.[7] The authors looked for differences in how the tumors developed, including the possible influence of perineal talc exposure. The only statistically significant association was with the use of "deodorizing powder," which was different from "baby powder" because the labels named "deodorizing substances and a variety of other free and bonded silicas" other than talc that were "potentially high in absestiform fibers." The authors were cautious about the implications for talc itself.

In 1997, Stella Chang and Harvey Risch published Perineal Talc Exposure and Risk of Ovarian Carcinoma, 79:12 Cancer 2396 (1997). It was a population-based case-control study in metropolitan Toronto, with 450 cases of borderline or invasive ovarian cancers and 564 controls. Controls were matched with cases by age group, and the analysis also considered as confounders the risk factors of oral contraceptive use, parity, breastfeeding, tubal ligation or hysterectomy, and family history of ovarian or breast cancer, which varied between the cases and controls as anticipated. The study found an "elevated" risk for both borderline and invasive ovarian cancer, but it was

[7] Borderline tumors are also called low-grade because they have low potential to become invasive and thus malignant.

statistically significant only for invasive cancer, and there was a marginally significant association with the duration of talc use, but not with frequency.

That study discussed two biological mechanisms, which had been postulated but not yet demonstrated, in which talc that migrated to the ovary could be a cause of ovarian cancer. One was talc's entrapment by inclusion cysts of ovarian epithelium during ovulation, while the other was talc's stimulation of entrapment of the surface epithelium, a phenomenon that had already been shown to be caused by "incessant ovulation." The authors observed that those mechanisms would be consistent with the author's own results, as well as with the results published in 1961 and 1971 by G.E. Egli and M. Newton in Transport of Carbon Particles in the Human Female Reproductive Tract, 12 Fertility & Sterility 151 (1961), and by W.J. Henderson and coauthors in 1971 in Talc and Carcinoma of the Ovary and Cervix, 78 J. Obstets. & Gyn. Br. Commw. 266 (1971), about finding talc particles in approximately seventy-five percent of examined ovarian tumors, and the results published in 1961 about the ability of nonmotile and inert carbon particles deposited in the vagina to migrate to the fallopian tubes.

In 1999, Cramer and coauthors published their population-based case-control study, Genital Talc Exposure and Risk of Ovarian Cancer, 81 Int'l J. Cancer 351 (1999). They noted the study subjects' age at first talc use and

their frequency and total years of use. The tumor subtypes of the cases were identified as serous, mucinous, endometrioid, clear cell, and other. They adjusted the results for age, parity, oral contraceptive use, obesity, family history of breast or ovarian cancer, tubal ligation, and the study location, which was eastern Massachusetts and New Hampshire. They found a statistically significant association of epithelial ovarian cancer with perineal talc exposure, whether by direct application or by transfer from talc applied to underwear or sanitary napkins. The association was most pronounced for invasive serous cancer and least pronounced for mucinous cancer. That study found a statistically significant dose-response trend when both cases and controls were considered together, but not when cases alone were considered. The study noted the difficulty of quantifying the amount of talc used in one application, and of correlating use to the times when the reproductive tract was open or closed.

The study further stated that the statistically significant association of talc use with ovarian cancer was consistent with the results of four other recent case-control studies, including Chang's. The nature of the results of that study and those other four, including the variation according to tumor histological subtype, suggested little confounding from recall bias, or from age, parity, or oral contraceptive use. It concluded that foreign body entrapment of talc

"appears able to induce histologic changes that are similar to those of asbestos, at least in the lungs," and that it was accordingly a plausible, although unestablished, mechanism of causation.

In 1999, Roberta Ness published a literature study, Possible Role of Ovarian Epithelial Inflammation in Ovarian Cancer, 91:17 J. Nat'l Cancer Inst., 1459 (1999). It was prompted by the observation that the hypotheses of causation by "excess" ovulation or by excessive gonadotropin and estrogen seemed to be incomplete explanations. Other studies suggested an association with epithelial inflammation, which could be caused by exposure to asbestos or talc, by endometriosis, or by pelvic inflammatory disease. Ness considered only epithelial tumors because they represented approximately ninety percent of all cases, and she did not distinguish between invasive and noninvasive tumors because they had similar risk factors.

The twelve epidemiologic studies of talc and ovarian cancer that she reviewed mostly found a significant association of perineal talc use with ovarian cancer, although some of them also found a dose response while others did not. She concluded that the consistent result of an association "in a series of well-conducted studies of varying design suggests" that talc use could "enhance" epithelial inflammation and thus promote cancer. However, Ness did not find any studies about the use of NSAIDs and ovarian cancer that

showed a statistically significant protective effect, or the lack of one, from their presumed anti-inflammatory effects. She concluded that much more study was needed to determine whether inflammation was a "central" element in ovarian cancer.

In 2000, Ness and coauthors published Factors Related to Inflammation of the Ovarian Epithelium and Risk of Ovarian Cancer, 11:2 Epidem. 111 (2000), a hospital-based case-control study of women diagnosed between 1994 and 1998 with borderline or invasive epithelial ovarian tumors. They found associations between ovarian cancer and several causes of inflammation, including talc use, as well as protective effects from agents like oral contraceptives that reduce inflammation. The association with talc use was statistically significant for all manner of direct use on the body, although when use on "genital/rectal and feet" was stratified by duration, the associations had somewhat weaker confidence intervals, and the association became statistically insignificant for one of the duration periods, namely, the period of five to nine years of such use.

Also in 2000, Dorota Gertig and coauthors published Prospective Study of Talc Use and Ovarian Cancer, 92:3 J. Nat'l Cancer Inst. 249 (2000). It used data from the Nurses' Health Study (NHS), a cohort study that was begun in 1976 with the enrollment of 121,700 female registered nurses in the United

States aged thirty to thirty-five years. In 1982, the subjects were asked to report whether they had ever used talc, whether they used it daily or weekly, and whether they used it perineally. A study cohort of 78,630 women was formed. Other factors, asked biennially, were oral contraceptive use, tubal ligation, and parity; family history of ovarian cancer was not asked until 1992. Additional questions addressed breastfeeding, age at menarche and menopause, and obesity.

From 1982 through June 1996, 307 cases of epithelial ovarian cancer were diagnosed in the study cohort. That study found a statistically significant association of 1.4, which it called a "modest elevation in risk," for ever users of talc and serous invasive ovarian cancer, but not for any other subtype of ovarian cancer. It further noted that the results "provide little support for any substantial association between perineal talc use and ovarian cancer risk." The study stated that tubal ligation did not affect the relative risk, which argued against the hypothesis that migration of talc through the fallopian tubes played a role in ovarian cancer, although it noted that the number of cases who had had tubal ligation was small.

The authors asserted that theirs was the first prospective study of talc use and ovarian cancer, and that being a prospective study eliminated recall bias and reduced selection bias. Conversely, they admitted the handicap of not

knowing the study cohort's ages at first talc use or their duration of talc use, which may have been a reason for the absence of a dose response. In addition, the "relatively short follow-up period may be inadequate to detect an association if the latency for development of ovarian cancer is more than 15 years."

In 2003, Michael Huncharek and coauthors published Perineal Application of Cosmetic Talc and Risk of Invasive Epithelial Ovarian Cancer: A Meta-analysis of 11,933 Subjects from Sixteen Observational Studies, 23 Anticancer Research 1995 (2003). It was a meta-analysis of sixteen observational studies about the association between ever perineal talc use and invasive epithelial ovarian cancer. The result was a statistically significant relative risk of 1.33.

However, the lack of a "clear" dose response prompted the authors to observe that the hospital-based studies showed a lower relative risk of 1.19 that was not statistically significant, while the population-based studies showed a higher relative risk of 1.38 that was statistically significant. They found that the difference suggested that the nominally stronger association for the latter reflected selection bias or uncontrolled confounding rather than a true risk.

In 2004, Paul Mills and coauthors published Perineal Talc Exposure and Epithelial Ovarian Cancer Risk in the Central Valley of California, 12 Int'l J. Cancer 458 (2004), a population-based case-control study of epithelial ovarian cancer that included questions about the frequency, duration, and particular years of perineal talc use. The odds ratio for ever users versus never users was 1.37 and statistically significant, but there was no dose response. The results differed by histological subtype, as in Gertig's study, and the highest odds ratio, 1.77 was for serous invasive tumors. The authors described the inflammation hypothesis as positing that inflammation produces oxidants that damage DNA, specifically the tumor suppressor genes, and that inflammation also reduces cytokine production with the possible result of altering cell growth and inhibiting apoptosis, which is the genetically regulated process by which a normal cell recognizes that it is damaged or senescent and proceeds to destroy itself. However, they noted the paucity of evidence to support the hypothesis as a cause of ovarian cancer.

In 2007, Cramer, John Godleski, and coauthors published Presence of Talc in Pelvic Lymph Nodes of a Woman With Ovarian Cancer and Long-Term Genital Exposure to Cosmetic Talc, 110:2:2 Obstets. & Gyn. 498 (2007), a case study of tissue samples, including lymph node samples, from a sixty-eight-year-old woman with serous ovarian cancer who had reported thirty

years of daily perineal talc use. Contamination from the study itself was ruled out as a source because the talc was found within macrophages in the tissue sample.

The authors stated that talc found in the lymph nodes supported new ways to think of talc's possible role in causing ovarian cancer. One would be inducement of an inflammatory reaction from deposition on the ovary. Another would be that chronic inflammation caused by talc in other parts of the reproductive tract, not just the ovaries, could cause a systemic decrease in the immune system's production of the antibodies to the MUC-1 protein whose overexpression is a feature of ovarian cancer.

Also in 2007, Amber Buz'Zard and Benjamin Lau published Pycogenol Reduces Talc-Induced Neoplastic Transformation in Human Ovarian Cell Cultures, 21 Phytotherapy Research 579 (2007), about their in vitro testing of a proprietary preparation of bioflavonoid derivatives of pine bark on ovarian tissue. They tested it on normal ovarian cells and nonepithelial ovarian tumor cells, as well as on polymorphonuclear neutrophils, a kind of immune system cell. They found that treating the cells just with talc increased the proliferation of precancerous cells, induced cellular transformations, and increased the generation of reactive oxygen species. All of those effects increased with length of exposure and dosage. However, when treatment of the cells with

their test preparation of the bioflavonoid derivatives preceded treatment of the cells with talc, their preparation "inhibited" the increase in cell proliferation, "decreased the number of transformed colonies," and decreased the generation of reactive oxygen species.[8] They concluded that the results "suggest that talc may contribute to ovarian neoplastic transformation."

In 2008, Hilde Langseth and coauthors published Perineal Use of Talc and Risk of Ovarian Cancer, 62 J. Epidem. & Cmty. Health 358 (2008), a pooled analysis of twenty case-control studies and one cohort study. They found that the fourteen population-based case-control studies showed an association of perineal talc use with ovarian cancer, of which ten were statistically significant, while the six hospital-based case-control studies showed associations that were not statistically significant. The cohort study showed no association. The cohort study and three of the four case-control studies that reported results by subtype gave "hints of higher risks of serous tumours related to talc exposure." While there was an overall association of talc use with ovarian cancer, the absence of an association in the cohort study and the absence of a "clear" dose response meant that the evidence to date was insufficient to "establish a causal association." However, the authors noted

[8] As related in the next part of this opinion, Cramer's report, and Omiecinski's report and testimony, explained the relevance of reactive oxygen species.

that the absence of a dose response could reflect "the crudeness of the exposure metric used," and they recommended additional studies with refined metrics, as well as better differentiation between talc products that contain asbestos and those that do not.

In 2009, Margaret Gates and coauthors published Risk Factors for Epithelial Ovarian Cancer by Histologic Subtype, 2010:171:1 Am. J. Epidem. 45 (2009), another prospective cohort study that relied on the NHS data. For a number of risk factors, they found that the factor's association with ovarian cancer varied according to whether the cancer's histological subtype was serous invasive, endometrioid, or mucinous, which may reflect the evidence that each subtype resembles a different kind of nondiseased tissue, or differences between the study populations in the distribution of cancer subtypes among the cases. In any event, talc use did not have a statistically significant association with any subtype.

In 2013, Kathryn Terry and coauthors published Genital Powder Use and Risk of Ovarian Cancer: A Pooled Analysis of 8,525 Cases and 9,859 Controls, 6:8 Cancer Prev. Research 811 (2013). Their analysis pooled the data from eight previous population-based case-control studies to estimate the association between lifetime talc exposure and ovarian cancer by histological subtype. There were 8525 cases of ovarian, fallopian, or peritoneal cancer and

9859 controls. Harmonization was needed for the data on the frequency and duration of genital talc use, but not for the data on the other potential risk factors or confounders, which included oral contraceptive use, parity, tubal ligation, obesity, age, race, and ethnicity.

The association of talc use with ovarian cancer was "stronger" for women who were obese than for those who were not, whereas there was no "significant" difference in the association for women who differed in parity or menopausal status, or in having endometriosis, tubal ligation, or a hysterectomy. There were likewise no differences in the association for women who started using talc after 1951, after 1961, or after 1971, although the association was somewhat lower but still statistically significant for those who started using talc earlier.

The study related that the histological subtypes of ovarian cancer were serous, endometrioid, mucinous, and clear cell; that tumors could be borderline or invasive; and that the most common subtype was serous invasive. Past studies showed that serous invasive had the strongest association with talc use. The authors noted that the only subtypes not showing a statistically significant association were mucinous borderline and mucinous invasive, which could have reflected either the relatively small number of tumors of those subtypes or some biological reason involving their molecular characteristics.

The study further reported that most of the increased risk appeared just from comparing "ever regular use to never use." The absence of a correlation between an increase in talc use and an increase in risk implied the absence of a dose response, which is considered an indicator of biologic plausibility. However, the lack of consistent evidence for dose response could also reflect "the difficulty inherent in accurate recollection of specific details of frequency and duration of genital-powder use," the different amounts of talc and other ingredients in various product formulations, or the possibility that "a modest exposure may be sufficient to increase cancer risk." Overall, the authors concluded that "genital powder use" was associated with a "small-to-moderate increase in risk of most histological subtypes of epithelial ovarian cancer."

In 2014, Serena Houghton and coauthors published Perineal Powder Use and Risk of Ovarian Cancer, 106:9 J. Nat'l Cancer Inst. dju208[9] (2014), a prospective cohort study that used data from the Women's Health Initiative cohort study (WHI). No statistically significant association was seen for ever use versus never use, or for increasing duration of use, even when stratified by age or tubal ligation status. However, the study had data only on the duration of use, not on frequency.

[9] This journal uses codes like "dju" and "djt" to locate articles, as the pagination of each article in this journal starts at 1.

Also in 2014, Britton Trabert, Ness, and coauthors published Aspirin, Nonaspirin Nonsteroidal Anti-inflammatory Drug, and Acetaminophen Use and Risk of Invasive Epithelial Ovarian Cancer: A Pooled Analysis in the Ovarian Cancer Association, 106:2 J. Nat'l Cancer Inst. djt431 (2014), a meta-analysis of population-based case-control studies. They concluded that aspirin had a statistically significant inverse relationship with invasive epithelial ovarian cancer, but that other NSAIDs and acetaminophen did not. The results were substantially similar for high-grade ovarian tumors of all histological subtypes, and also for borderline serous tumors. They considered their results to be general rather than specific support for the hypothesis that inflammation played a role in ovarian cancer, because "[t]he pharmacological effects of NSAIDs that lead to reduced risks of cancer or improve cancer prognosis are not well understood and may differ by cancer site."

Later in 2014, Trabert and coauthors (not including Ness) published Pre-diagnostic Serum Levels of Inflammation Markers and Risk of Ovarian Cancer in the Prostate, Lung, Colorectal and Ovarian Cancer (PLCO) Screening Trial, 135:2 Gynec. Oncol. 297 (2014). It was a prospective case-control study that took advantage of the collection of blood samples from participants in a screening trial for those four kinds of cancer to look for an association between the level of numerous chemical markers of inflammation and a subsequent

increased risk of ovarian cancer. The authors discussed human and animal studies that provided evidence of how both inflammation connected to ovulation and other inflammatory processes may play a role in ovarian cancer, including the possibility that some ovarian cancers, notably the subtype of serous invasive, could arise from inflammation of the fallopian tubes or of endometriotic lesions as well as of the ovaries themselves.

After statistical analysis to correct for the influence of obesity, parity, hormone therapy, oral contraceptive use, aspirin or ibuprofen use, and family history of ovarian or breast cancer, the authors reported evidence of an association with ovarian cancer that was statistically significant for two markers and equivocal for several others. They saw the study as having limited power to detect associations for most subtypes of ovarian cancer, but as yielding "compelling" evidence of an association between several inflammation markers and serous ovarian cancer. Some of the inflammation markers were associated with other cancers, so they noted the need for additional research to identify particular markers with particular cancers, and to correlate the level of such markers in the blood with their level at the sites where inflammation could lead to ovarian cancer.

The record contains the abstract of Does Talc Exposure Cause Ovarian Cancer?, 25 Int'l J. Gyn. Cancer 51 (2015), which Ness published in 2015.

The abstract called the underlying study a "formal systematic analysis of talc use and ovarian cancer," based on numerous case-control and cohort epidemiological studies, meta-analyses, and "basic science studies," which were "reviewed and graded for quality." Ness conducted analyses on the data in the aggregate and also by histological subtype, in line with the Hill factors. She concluded that those studies "suggest that talc use causes ovarian cancer," because "almost all [of the] well-designed studies" showed that talc use increased the risk of ovarian cancer by thirty to eighty percent, which she distilled to an "attributable risk" of twenty-nine percent. The association was "more specific" for serous ovarian cancer. She noted that the studies that addressed dose response found it to exist for both duration and frequency of exposure.

The abstract stated that systematic bias could be "excluded" because the nature of the studies minimized recall and selection bias, and because they conducted multiple assessments of other risk factors for ovarian cancer. It declared inflammation to be "a plausible biological mechanism" because it was "known to cause other epithelial cancers."

In 2016, Cramer and coauthors published The Association Between Talc Use and Ovarian Cancer, A Retrospective Case-Control Study in Two US States, 27:3 Epidem. 334 (2016), about the population-based case-control

study mentioned above of 2041 cases and 2100 controls in eastern Massachusetts and New Hampshire.[10] That study had three consecutive five-year enrollment periods between 1992 and 2008, and this study purported to be the first to address the data from all three periods. The 1999 Cramer study had addressed only data from the first period, while the 2008 Gates study combined data from the second period with NHS data, and the 2013 Terry study combined data from the third period with data from several other studies.

The authors noted that the subjects reported age at first use, years of use, uses per month, and whether the application was perineal, on another body area, or on an item that touched the body. Only perineal use, either alone or with additional forms of use, had an odds ratio greater than 1.0 for epithelial ovarian cancer, and it was statistically significant. For those users, the overall results were the statistically significant odds ratio of 1.33, with a trend of increasing risk for increased frequency of talc use, but not for increased duration. For cases with more than twenty-four years of perineal use, the association was stronger for the histological subtypes of borderline serous, borderline mucinous, invasive serous, and invasive endometrioid.

---

[10] The record contains the 2015 prepublication version. The published version, which is no different, is available at https://www.researchgate.net/publication/5512175_Perineal_use_of_talc_and_risk_of_ovarian_cancer. The 2015 version is the one that Cramer cited in his expert report.

While the genital talc users were more likely to be older, heavier, asthmatic, and regular users of analgesics, sensitivity analysis by logistic regression and other methods showed that none of those factors was a confounder. The authors applied what they called the convention of regarding a factor to be a confounder only if adjusting for it changes the odds ratio by ten percent in either direction.

The authors called their results consistent with the 2013 Terry pooled analysis. They addressed the possibility of recall bias by applying a sensitivity analysis. In the absence of external records to verify the study subjects' reported use or nonuse of talc, which they would have used to perform that analysis, they used a surrogate analysis, namely, the sensitivity analysis of alcohol use in the NHS evaluation of alcohol use and breast cancer, in which retrospective recall could be compared to verifiable prospective data. The rate of accurate recall was found to have been ninety-one percent, meaning a nine percent misclassification rate. The authors noted that twice as much misclassification of talc use, or a rate of eighteen percent, would have been required for their observed odds ratio to lose statistical significance. They then discussed several reasons that made their odds ratio less likely than that to result from recall bias. Those reasons were the greater likelihood of accurate recall of ever using talc as opposed to remembering the specific degree of use,

and the tendency of recent studies to show lower odds ratios than older studies did, notwithstanding the increase over time in publicity about the possible association of talc with ovarian cancer.

The authors of that 2016 study found that the dose response was "more apparent" for cases who were premenopausal or who were "heavier or postmenopausal users" of hormone-replacement therapy when diagnosed. Other factors in premenopausal women, including weight, breastfeeding, and alcohol use, may also have been "effect modifiers" rather than just confounders because they tended to alter estrogen levels, which "may have multiple effects on immune cells," such as causing macrophages to scavenge particulates like talc that they would otherwise disregard. Those women comprised the categories that showed more of a dose response, so the possibility that those factors had multiple effects that might make the immune system overly responsive to talc, combined with the documented ability of talc to migrate to the upper reproductive tract, suggested that "a framework" existed for positing a mechanism "involving chronic inflammation" by which talc at least promoted ovarian cancer.

The authors acknowledged the novelty of finding no association between ovarian cancer and perineal talc use by postmenopausal women who were not receiving hormone replacement therapy. However, the WHI study, which

enrolled only postmenopausal women, concluded that hormone replacement therapy was just a confounder, not an effect modifier. The authors of that study did not see the WHI study as disproving the possibility that altered estrogen levels could be an event modifier in premenopausal women, so they did not see it as discrediting their suggestion that the combined agency of altered estrogen levels and talc use could cause chronic inflammation that facilitated the development of ovarian cancer.

IV.

We now turn to the discussion of plaintiffs' experts, their reports, and testimony.

A. Daniel Cramer. At the time of the Rule 104 hearing, Cramer was a professor of obstetrics, gynecology, and reproductive biology at Harvard Medical School, as well as a professor of epidemiology at Harvard's T.H. Chan School of Public Health. He headed a research division of obstetric and gynecological epidemiology with a particular focus on ovarian cancer. He had performed epidemiological research for more than thirty years, co-authored many published scientific articles on environmental and genetic causes of ovarian cancer, authored several chapters in books on oncology and epidemiology, and authored or co-authored several publicly presented abstracts on epidemiological studies of ovarian cancer.

Cramer's February 1, 2016, expert report on general and specific causation for Carl cited 101 published studies, including his own earlier studies. One was a 1996 study by Debra Heller and others that found cancerous human ovarian and uterine tissue samples to contain "birefringent" particles that could have been talc.[11] Cramer's own 1982 epidemiological study, a population-based case-control study, was the first to find a statistically significant association between perineal talc use and epithelial ovarian cancer. Cramer cited twenty-five additional published studies through 2014 of talc and ovarian cancer; all of them found an association, and in twelve of them the risk was statistically significant.

Cramer also cited two meta-analyses, by Gates in 2008 and Terry in 2013, of previously published data that found a significantly increased risk for ovarian cancer from talc use. He explained that a meta-analysis was "more powerful" and provided "a more precise estimate of the association" because the ninety-five percent confidence interval was narrower for that combined assessment than in the underlying studies individually.

Cramer was aware of five meta-analyses on talc and ovarian cancer, including his own from 1999, Huncharek's in 2003, and Langseth's in 2008 in

[11] Debra Heller and others, The Relationship Between Perineal Cosmetic Talc Usage and Ovarian Talc Particle Burden, 174:5 Am. J. Obstets. & Gyn. 1507 (1996).

connection with an IARC review. He described the studies that each one incorporated and related that each of those meta-analyses found a statistically significant association. He also described minor issues in some of the underlying studies concerning the distinction between perineal use of talc and other uses, or combined use, before opining that adjusting the odds ratios in those studies to conform better to a model comparing subjects who were perineal ever users versus perineal never users would have had little effect on any of the results.

In addition, Cramer performed a new meta-analysis on the entire body of data in the studies and meta-analyses that he had related. There was no significant heterogeneity among them, even though two of the studies were cohort studies while the others were case-control studies. The "summary" odds ratios for the risk of ovarian cancer between ever use subjects and never use subjects was 1.29, and it was statistically significant.

Cramer then discussed the Hill factors for an association to support an inference of causation. He opined that the result of a statistically significant association was consistent in studies in the United States, Canada, England, China, and Australia, which established geographical and ethnic diversity of the study populations. The results were also consistent between the case-

control studies that were hospital-based and population-based, and there was no significant heterogeneity among them.

Cramer noted that some researchers had called the tendency of cohort studies to report a lower odds ratio for talc and ovarian cancer than in the case-control studies a sign that the case-control studies had recall or reporting shortcomings. He disagreed, on the ground that one would expect more recall or reporting bias in the more recent studies, due to increased publicity about the potential link between talc use and ovarian cancer, yet the odds ratios in the recent case-control studies were not higher than in the earlier ones. He believed instead that neither cases nor controls were likely to be inaccurate about "daily or weekly use of talc carried on for decades[,] which is where the risk for ovarian cancer from talc use lies."

Cramer also mentioned selection bias, which he described as the possibility that the exposure history of the cases or the controls was not representative of the portion of the general population that the study intended to address. He explained that "significant correlations" in the reported response rates between cases and controls would suggest selection bias, and that his 2016 meta-analysis did not find any.

Cramer noted that confounding can occur in both case-control and cohort studies. He observed that most talc studies adjusted for age and known risk

factors, including parity and oral contraceptive use. Some studies, including his own 2016 study, had odds ratios that remained significant after adjustment for obesity. Indeed, Cramer's latest study did not find that obesity or any of twenty-three other potential confounders changed the crude odds ratio by as much as ten percent, the conventional threshold for a confounder. As additional confirmation, Cramer cited a study, published by John Whysner in 2000, as finding no evidence that potential confounders increased the risk of ovarian cancer for women who had used cornstarch instead of talc.

As for the strength of the association, Cramer explained that Hill stated that an odds ratio of less than 2.0 can be strong enough to indicate causality as long as the association did not arise from bias, confounding, or random error. Cramer cited genome association studies that were analogous to the meta-analyses of talc and ovarian cancer in the number and heterogeneity of study subjects, and he stated that their authors inferred causation on statistical results comparable to those in his own studies. On those bases, he opined that an odds ratio of 1.3 was strong enough to support an inference of causation.

Cramer opined that questions about dose response required information about the frequency and duration of talc use. He acknowledged the difficulties arising from the lack of a standard measure for the amount of talc used in a perineal application, the amount entering the body, and the amount reaching

the upper reproductive tract. Nonetheless, "larger and more recent case control studies" that he cited, including his own from 2015, showed a dose response according to the estimated number of applications, especially when the analysis was limited to users, or to subjects whose upper reproductive tracts were open to particulate transmission.

For biologic credibility, Cramer stated that the association must "make[] sense in terms of what is known about the biology of the cancer" and about whether animal or cell-line experiments "support an association." He cited several studies as proving that talc particles can migrate as far as the ovaries. After describing the theory in his first paper that talc particles can "cause changes predisposing to ovarian cancer," he cited Buz'Zard's 2007 study for its finding that talc-induced changes in ovarian cell proliferation that were "indicative of malignancy" could be increased by anti-inflammatory agents, and he noted that the finding suggested "a role" in ovarian cancer for the reactive oxygen species that are part of the response when inflammation stimulates the immune system into action.

Cramer's most recent theory relied on a model in which chronic inflammation in the upper reproductive tract blunted the immune system's production of the antibodies that respond to the class of cellular-surface proteins called mucins, which include the molecular markers of ovarian

cancer. When the immune system is functioning normally, it produces protective antibodies when those mucins are "over-expressed," which occurs "during inflammatory, infectious, hormonal, or neoplastic events." He called ovarian cancer "a mucin secreting cancer," and he opined that the data from case-control and cohort studies showed that increased levels of anti-mucin antibodies were associated with decreased risks of ovarian cancer, while decreased levels of those antibodies were associated with increased risks of ovarian cancer.

Cramer further explained that women with ovarian cancer and long-term talc use had blood-test results before the start of cancer treatment that indicated chronic inflammation. He then opined that long-term talc use could cause chronic inflammation in pelvic lymph nodes, that the immune system's response to such chronic inflammation would eventually fatigue it, and that the fatigue would blunt the immune response to the over-production of mucin in the ovaries and allow cancer to develop.

Cramer noted the 2014 statement by the NCI that the results of WHI and NHS did not support an inference of causation for talc and ovarian cancer. He observed that WHI enrolled only women of an average age well past that of menopause, a population that had a lower association between talc use and ovarian cancer than for premenopausal women, and that it failed to identify

cohort members who had their ovaries removed during the study period. For NHS, talc exposure was assessed only upon enrollment and was assumed to remain constant during the twelve-year study period, despite the likelihood that nurses would have been aware of the "considerable publicity" about talc and ovarian cancer and might have reduced their talc use in response.

Cramer opined that Carl's obesity, nulliparity, and reported frequency and duration of perineal talc use were the "major factors that could have contributed to" her ovarian cancer, which was a serous borderline tumor. Carl's reported talc use amounted to an estimated 5980 applications over twenty-three years, and Cramer's analysis of "data supplied to the Defense" in an out-of-state case about perineal talc use as a cause of ovarian cancer yielded a statistically significant odds ratio of 2.05 for serous borderline tumors in women with more than 5040 applications. He performed a meta-analysis of studies about obesity like Carl's and ovarian cancer, and another meta-analysis of studies about parity and ovarian cancer. The odds ratios that he calculated were lower than that for talc use like Carl's and ovarian cancer, so he opined that her talc use was "more likely than not . . . the major cause" of her cancer.

Cramer explained that Carl had a "very low likelihood" of the BRCA mutation that can increase the risk of ovarian cancer, based on the absence of a family history of ovarian cancer and on a study in Ontario from 2001 in which

none of the cases with borderline ovarian cancer had that mutation. She had used psychotropic medication, been employed as a hairdresser for seven years, and had a smoking history. The ovarian cancer studies that addressed those potential risk factors were inconsistent, failing to show a statistically significant association with serous borderline ovarian cancer.

Cramer issued his expert report on general and specific causation for Balderrama on February 23, 2016. The opinions and explanations on general causation were the same as in his report for Carl. Balderrama was thirty-six years old when he issued this report, she had no children, she had never smoked, and she was obese. Multiple examinations starting in October 2011 to assess her infertility ended with surgery in November 2012 that included removal of her ovaries. Pathology revealed an endometrioid tumor of the right ovary and an endometrioid invasive tumor of the uterus.

The pathologist could not determine whether the tumors were related. Cramer's colleague, Dr. William Welch, an expert in gynecological pathology, reviewed pathology slides and concluded that the tumors were independent primary tumors. Cramer agreed, based on studies showing that it was relatively rare for an ovarian endometrioid tumor to be the secondary manifestation of another endometrial neoplasia. Cramer explained that the

primacy of the ovarian tumor allowed him to analyze and weight its risk factors separately from such an analysis and weighting for the uterine tumor.

Cramer cited four studies on obesity and endometrial cancer to opine that the association of obesity with ovarian cancer varied by histological subtype. He performed a meta-analysis of the eight studies that reported odds ratios for ever use of talc and endometrioid ovarian cancer versus never use, and he found a statistically significant summary odds ratio of 1.4. Only a small number of cases among those studies were premenopausal like Balderrama and reported talc use that approached her estimated 9700 applications, so he used the data from an out-of-state litigation "for all endometrioid cases" of ovarian cancer, apparently meaning premenopausal and menopausal, categorized by number of applications. His result for cases who had more than 6000 applications and were obese was a statistically significant odds ratio of 1.79.

For the effect of parity, Cramer found five studies and performed a meta-analysis that yielded a statistically significant summary odds ratio of 1.60. Balderrama reported having used oral contraceptives to regulate her menstruation, but her lack of recall about the duration of such use and the irregularity of her cycle made it impossible to determine whether that use might have conferred any degree of the known protective effect against

ovarian cancer through the suppression of ovulation and thus of its attendant inflammation.

Cramer cited two studies that reported results about the risk that Balderrama's degree of obesity posed for endometrioid ovarian cancer. One reported a statistically significant odds ratio of 1.86 compared to nonobese study subjects, the other an odds ratio of 1.2 that was not statistically significant. He opined that the odds ratio for talc was higher than the combined "inconsistent" odds ratios for obesity, which made Balderrama's talc use "more likely than not" the "major cause" of her endometrioid ovarian cancer. He added without elaboration that Godleski's finding of talc in Balderrama's ovarian tissue was a factor in his opinion. By contrast, when Cramer performed that analysis for Balderrama's independent uterine tumor, he determined that its primary cause was her obesity rather than her talc use, even though the association of obesity with that tumor's histological subtype was much lower than the association of obesity with uterine cancer in general.

Cramer testified that potential confounders must at least be named, not just presumed as in some industry criticism of certain studies. He added that no scientist had declared an odds ratio of 2.0 to be the threshold below which causation may not be inferred. He criticized the NCI's statement of no association between talc use and ovarian cancer by explaining what he saw as

its overreading of certain studies that it cited, and for citing only four studies when the literature contained more than twenty-five. He criticized the FDA's April 2014 letter on talc and ovarian cancer for failing to cite any authority when it declared the lack of causality.

For Colditz's statement on a hospital website that an association of 1.1 to 1.5 is a "weak" risk, Cramer called it necessarily reductive so that patients could understand it, and that it was neither Colditz's nor anyone else's idea of a scientific statement. Responding to an objection by a reviewer of his 2016 study about his "dicing and slicing" the data in order to explain away confounders, Cramer said that the objection was invalid because such data analysis is exactly how one tests for confounders.

Cramer explained that cohort studies must track their subjects during their entire duration for both age and cumulative exposure at each data-collection interval, or they may risk reporting an injury rate that looks steady across the intervals, and miss the true rate if the injury is one that develops more slowly than expected. More generally, what mattered in a cohort study was not so much the size of the study population as "the number of cases found and the quality of the exposure data that the cohort started with." It was an increase in the number of cases, not in overall study population, that would afford a "more precise" odds ratio and a narrower confidence interval.

Cramer said that Balderrama developed her tumors early enough to raise the question of genetics generally, but that nothing in her family history of cancer stood out as suggesting a genetic cause. Cramer acknowledged that her relative risk of 1.86 for ovarian cancer from obesity represented a significant risk that she could have developed endometrioid cancer from that cause alone.

Cramer then explained that the quartiles for talc exposure in his analysis for Carl were different than in his analysis for Balderrama because their exposure periods were different, but that the quartiles still yielded a reasonable set of exposure categories. He used the literature to estimate Carl's relative risk for ovarian cancer from obesity at 1.75, but he did not stratify the data in that estimate by degree of obesity, even though she was not much less obese than Balderrama.

B. Graham Colditz. Graham Colditz testified as an expert epidemiologist specializing in identifying avoidable cancer risk factors. He was licensed to practice medicine in Australia, held a doctoral degree in epidemiology and public health, was a professor at Washington University School of Medicine, and the associate director for prevention and control at Siteman Cancer Center, an NCI-funded comprehensive cancer center.

Colditz issued his expert report on general causation on July 31, 2015, which cited sixty-three published studies. On "the totality of all evidence and

the continuing accrual of new studies," he opined that genital talc use "can cause ovarian cancer." He did not address specific causation.

Colditz noted that Hill provided a framework for addressing the issues in "summarizing evidence," which included strength of association, consistency of studies in finding an association, temporality, dose response, biologic plausibility, "coherence," "experimentation," and "analogy." For the association of talc exposure to ovarian cancer, Colditz identified the "key" issues as consistency of association, dose response, and biological plausibility.

Colditz described his methodology as starting with "a systematic search and review of the literature" including his own prior research, analyzing "experimental, clinical and epidemiological studies and data," and applying his "skills in research synthesis." He then assessed the epidemiological studies for potential biases and confounding, and observed that some meta-analyses paid "insufficient attention to the quality of the exposure and outcome measures" in the underlying studies.

Colditz summarized the grounding for his opinions as the epidemiological studies that "show" an increased risk of ovarian cancer from talc use and "support" a dose response. His basis for believing talc to be a biologically plausible cause of ovarian cancer was that "[t]alc can travel to the

ovaries causing an inflammatory response" and that "the inflammatory mechanism is consistent with the increase in risk of ovarian cancer."

Colditz related that most studies of talc and ovarian cancer were case-control studies, most were population-based, and focused on "detailed assessment of exposure among cases and control subjects." For the epidemiological studies published in 2006 or earlier, Colditz relied on the summaries of their evidence in a 2006 IARC report not included in this record, which summarized the epidemiological studies to that date, the evidence from in vitro studies, and "other sources of evidence."

Colditz described the IARC 2006 report, the 2008 Langseth study, and a 2006 study by Robert Baan as concluding that talc was "a possible carcinogen." He stated that the population-based case-control studies showed a statistically significant association of 1.4 between ever use and ovarian cancer. He added that in a part of the IARC study "[f]ocusing on [eight] higher quality studies," which included five of the studies in this record (Cramer 1982, Chang 1997, Cramer 1999, Ness 2000, and Mills 2004), the IARC found that the rate of perineal talc use among controls ranged from sixteen to fifty-two percent, and that the relative risk of ovarian cancer correspondingly increased from 1.30 to 1.61. Furthermore, four of the five

studies that reported results by histological subtype suggested that talc exposure created a higher risk of serous tumors than of other subtypes.

Colditz explained that the WHI study participants were at an average of ten years after menopause upon enrollment, and that the talc users were asked to report duration, but not frequency or whether their use was current. The study assumed no changes in a participant's status during the 12.4-year study period, including no surgical removal of an ovary. The study reported no association between talc use and ovarian cancer, but Colditz saw "considerable" limitations in the data that it collected and the ensuing analysis. For NHS, the cohort was thirty to fifty-five years old at enrollment, yet talc use was similarly determined at enrollment by only one parameter, in that case frequency instead of duration, and it was assumed to remain constant.

Colditz cited Gertig's 2000 study as the first analysis of NHS data. No association was found for ever users without regard to subtype, but when subtype was considered, a "significant increase in risk" appeared for invasive serous cancer. Colditz then cited the Gates's 2008 study as finding a significant increase in risk from "regular talc use," with the risk being "somewhat stronger" for invasive serous cancer than for ovarian cancer overall without regard to subtype.

Colditz explained that Terry's 2013 study had "the strongest analytic approach," because it did not just combine the reported results of individual summaries, but rather obtained all the data and used "common definitions and analytic methods" to analyze the data for each individual patient, which reduced the potential bias from differences in methodology. That approach was applied to the data from eight case-control studies, some of which were updated to include additional cases and controls since their publication, for a total of 8525 cases and 9859 controls. The analysis controlled for the established risk factors for ovarian cancer, which included age, parity, oral contraceptive use, tubal ligation, obesity, and race and ethnicity. Colditz called the statistically significant association of 1.24 for genital talc use and ovarian cancer compared to never use a "modest increase in risk." The risk was higher for "cancers defined by cell subtype" and for borderline serous tumors.

Colditz recognized that Terry's 2013 study found a dose response only for non-mucinous tumors, and only when the entire study population was considered, with no dose response when only users were considered. However, four other studies showed a significant dose response, and three of them were among what the IARC called the eight higher quality studies.

On magnitude of risk, Colditz insisted that it not need reach 2.0 to support an inference of causation. He explained how the IARC had classified a combined hormone therapy as a cause of breast cancer based on WHI data that showed the relative risk to be from 1.24 to 1.26.

Colditz opined that "the quality and depth of exposure assessment" were fundamental questions in evaluating an epidemiological study. He opined that case-control studies may have more complete assessments of an exposure if that is their sole or primary focus, whereas cohort studies "typically relate lifestyle exposures to a broad range of conditions" and have less room in their questionnaires for stratification questions at enrollment or for follow-up questions about changes in status. The point was not that one kind of study was better or more reliable, but rather that "the details of exposure assessment" at enrollment and over time were important.

Colditz discussed biological plausibility briefly, by citing the 1999 Ness study, a 2009 study published by Jack Cuzick and coauthors that is not in the record,[12] the 2014 Trabert study, and the 2014 Trabert and Ness study. He believed that they "established that talc can travel to the ovary, it causes an

[12] Jack Cuzick, Aspirin and Non-Steroidal Anti-Inflammatory Drugs for Cancer Prevention: An International Consensus Statement, 10 Oncol. 501 (2009).

inflammatory response, and this mechanism is consistent with the increase in risk of ovarian cancer that is observed."

In his testimony, Colditz opined about ovarian cancer in general, without specifically discussing different subtypes. He repeated the descriptions of epidemiological studies, meta-analyses, and the primacy of study design to reliability that were in his report. He also repeated his report's description of his methodology, and of his views on the typical limitations of cohort studies, using NHS as an example.

Colditz believed that the IARC's 2006 review of talc and ovarian cancer was "full and complete," at least for its time. He added that the successive meta-analyses, each to some degree expanding upon its predecessors, gave a sense of the accumulating evidence of talc's association with ovarian cancer. He thought that Cramer's 2016 study truly minimized confounding. On the totality of the evidence, Colditz opined that talc use causes ovarian cancer.

Colditz agreed that the cohort studies and the hospital-based case-control studies did not report a statistically significant association between talc use and ovarian cancer, and that the population-based case-control studies had mixed results. He criticized hospital-based case-control studies for uncertainty about their "catchments" for different diseases, presumably meaning that the study populations may have additional diseases that are confounders for the

disease being studied. He asserted that the NCI was funding population-based case-control studies rather than hospital-based ones for that reason, and that case-control studies intended for publication in peer-reviewed articles will similarly attract funding only if they are population-based.

Colditz declared that a risk ratio did not have to exceed 2.0 to be meaningful, and he added that in comparing study results, a lower relative risk may be more meaningful if it comes from a larger study, for which size alone often affords a tighter confidence interval. For those reasons, calling a study weak or strong based solely on the relative risk ratio that it generated would be unsound.

Colditz acknowledged that, while his report cited studies supporting acceptance of inflammation as a plausible mechanism, it did not cite studies or other literature on the plausibility of talc migration to the ovary. When asked to address migration further, he responded that "I believe others have written reports and detailed on that."

The trial judge asked Colditz to elaborate on the theory about inclusion cysts in Cramer's 1982 study, and he responded by describing the theory as postulating that when an ovary's surface epithelium is disrupted by ovulation, the immune system treats it as an inflammatory event, with talc that is present on the surface getting entrapped in the inclusion cyst during the repair of the

ovarian surface. When the court asked Colditz if he had found any other peer-reviewed articles in which that theory had been discussed, he replied that he did not know of one that discussed inclusion cysts, and that there was a need for "continuing studies to understand this whole process better."

C. John Godleski. John Godleski was at the time of the hearing a Harvard Medical School professor of pathology. He had published numerous papers on electron microscopy and environmental pathology. He conducted a pathology research group, and he was an expert in diagnosing foreign material in all body tissues.

Godleski analyzed tissue samples from Carl and Balderrama. For Carl, he used the samples to confirm the diagnosis of serous borderline cancers in the right and left ovaries with metastases to two lymph nodes. The pathology report from the hospital that supplied the samples stated that Carl also had "invasive tumor implants" on her uterus and elsewhere within her peritoneum.

Godleski's report described how his laboratory observed its protocols to avoid contaminating the tissue samples. The laboratory then used polarized light, followed by a scanning electron microscope with an energy dispersive X-ray analysis system, to identify birefringent particles in one ovary and one lymph node. Spectral analysis showed that most particles were of kinds normally present, while some other particles contained magnesium, silicon,

and oxygen "in the proportions expected with talc," which was enough to identify them as talc.

The report explained that the testing used "an extremely small volume of tissue," and that the number of talc particles indicated that "substantial amounts of talc were present in this patient," including "within the ovary/tumor and draining lymph nodes," which was consistent with Cramer's published finding about one ovarian cancer patient who had "large amounts of talc . . . in lymph nodes draining the pelvis." Godleski concluded that "the talc found in this case" was "evidence for a causal link between the presence of talc and the development of" Carl's ovarian cancer.

For Balderrama, Godleski's report related the use of similar procedures to distinguish particles normally present from particles with the composition of talc in her right ovary, endocervix, uterine wall, and some lymph nodes. In similar fashion, the report explained that substantial amounts of talc were present in Balderrama, and reached the same conclusion, which was that the talc was "evidence for a causal link" between the talc's presence and the occurrence of ovarian cancer.

Godleski testified that his belief in a possible causal link between the talc particles that he found in Carl's and Balderrama's tissue samples and their cases of ovarian cancer was based simply on the consistency of his findings

with the reports in some epidemiological studies of a causal link between the presence of talc and ovarian cancer. The presence of talc just "add[s] evidence to the epidemiologic story," and Godleski did not presume to proffer evidence of biologic causation himself, other than to state that he believed the talc was present because it had been collected by macrophages. Indeed, he had no reason to doubt the findings of Carl's and Balderrama's treating pathologists that neither of them had a "talc-related inflammatory reaction."

D. Curtis Omiecinski. Curtis Omiecinski, who had a Ph.D. in pharmacology, was a professor of molecular toxicology at Penn State University. His discipline required study in chemistry, biochemistry, biology, physiology, molecular biology, and genetics and in how they "come together." His main work was to "make predictions about the interactions of chemicals [and] environmental agents on disease status and human health in particular." Plaintiffs submitted a report that Omiecinski had issued in April 2015 in an out-of-state litigation on talc and ovarian cancer.

Omiecinski's report stated that "particulate exposures in general often evoke inflammatory responses within the affected tissues and organs." Inflammation and its "pathways" have been "recognized" as part of the cause of prostate cancer, and they are "likely" part of the cause of "epithelial ovarian cancer" as well. In general terms, when particles cause inflammation,

macrophages detect and engulf them and release chemokines, which recruit leukocytes and facilitate their entry into cells, prompting the cells to generate reactive oxygen species that can incidentally damage genetic material in ways that lead to mutations. Mutations, and also the cell proliferation that inflammation promotes, contribute to the early stages of cancer, which develops through multiple stages.

The observation of several factors that are present when inflammation and ovarian cancer are also present has inspired hypotheses about inflammation as a cause of cancer. However, while much of carcinogenesis is common to all cancers, the differences among normal tissue types in sensitivities and in the ability to repair genetic damage or force the death of abnormal cells may also exist for the corresponding variety of "tissue-selective cancers" that differ at least partially in their molecular pathways.

Omiecinski cited "[s]everal lines of evidence" showing that particulates like talc can migrate from the perineum to the upper reproductive tract. He also cited in vitro studies, including Buz'Zard's, of the response of cultured human cells to inflammation and the oxidative stress that it creates. On that basis, he opined that talc in certain situations can "trigger" inflammatory responses that cause the creation of reactive oxygen species. Although he was not an epidemiologist, he believed that the weight of the corpus of

epidemiological studies of talc and ovarian cancer demonstrated enough associations to support his opinion that chronic perineal exposures to talc were "predisposing and causative contributors" to the development of epithelial ovarian cancer.

In his testimony, Omiecinski restated his opinion that perineal talc can migrate to the ovaries, that talc in ovarian tissue can cause inflammation, and that such inflammation can "initiate" cancer. He developed his opinion by reviewing the literature. His search yielded seventy-one peer-reviewed articles, including approximately three dozen epidemiologic studies that reached varying conclusions about the association of talc with ovarian cancer. He focused on the biology and genesis of ovarian cancer, the migration of particles through the reproductive tract to the ovaries, the differences between talc and other particles, the cellular effects of talc exposure, and possible mechanisms for chronic talc exposure to cause ovarian cancer. He also looked at websites including those of the IARC, the NCI, and the FDA.

Omiecinski explained that one of the cellular effects of inflammation is the process that leads to the generation of reactive oxygen species, which could then initiate a process leading to cancer. Those oxygen species can be beneficial by killing infection cells, but when inflammation is not caused by infection, they can instead act upon and damage the DNA of healthy cells, and

the mutated DNA can initiate carcinogenesis by signaling those cells to proliferate.

Omiecinski observed that in vitro studies were valuable because they permitted observation of "live cellular systems" in precisely controlled conditions. There were in vitro studies on many different particles in addition to talc, and he opined that they were similar in showing an inflammatory response that could "be manifested in increased proliferation ability" of the damaged cells.

Omiecinski noted that Buz'Zard's in vitro study, about the effect of talc on granulosa ovarian cells and on epithelial cells, had three results characteristic of the progression toward cancer. They were the increase in reactive oxygen species; the increased rates of cell proliferation that are evocative of cancer's uncontrolled proliferation; and the increase in cellular "neoplastic transformation" and "dedifferentiating," which meant departures from the cell's proper morphology and functioning toward the aberrance that typifies cancer cells.

Omiecinski agreed that his opinion and explanations were not inconsistent with the proposition that reactive oxygen species that arise solely from inflammation may cause cellular damage that leads to cancer. He then

agreed with the coherence of a theory that the monthly inflammation due to ovulation may be enough to initiate that process.

V.

We have provided exhaustive details of the reports to support our conclusion that plaintiffs' experts provided admissible opinions meeting the Manual and Hill protocols. They relied upon significant studies that the relevant scientific field accepted as suitable for such reliance. The reasons that Cramer and Colditz gave for finding certain epidemiological studies more pertinent than others did not conflict with the scientific community's principles for interpreting and relying upon studies. They neither misread or misrepresented study results, nor relied on studies that represented less than a substantial portion of the available scientific literature. They anchored their opinions on the studies regarding biologically plausible mechanisms that even governmental and agency resources recognized as plausible.

Although the Manual observed that larger study populations, where possible, were more reliable, the Manual also acknowledged that size alone was not a paramount foundation for reliability. It did not declare cohort studies inherently more reliable than case-control studies due to population size or any other design element.

Cramer's explanation of how he interpreted and relied on epidemiological studies was consistent with the Manual and Hill. He disagreed with the view of some researchers that the lower odds ratios reported in the cohort studies exposed the presence of recall or reporting bias in the case-control studies. He explained that study subjects were unlikely not to remember the decades-long use of talc on a daily or weekly basis that he said was needed for talc to become a risk factor, and that the absence of such bias was demonstrated by the consistency over the years in the odds ratios from case-control studies, notwithstanding the growing publicity about the suspected association of talc with ovarian cancer. Cramer further explained how he tested for selection bias in his 2016 case-control study and did not find any. He added that cohort studies must repeatedly obtain data about their participants' cumulative exposure, in order to detect the true association if the disease's latency is greater than expected.

Cramer then noted that confounding can occur in any study, that most studies addressed age and known risk factors, and that the testing for confounders in his 2016 study found their influence to be too small to affect the results. He also explained that the authors of genome association studies that were analogous to meta-analyses of talc and ovarian cancer in the number and heterogeneity of study subjects inferred causation upon statistical results

comparable to those in his own studies. For all of those reasons, Cramer opined that an odds ratio of 1.3 was strong enough to support an inference of causation.

Colditz's explanation of his reliance on studies was likewise consistent with the Manual and Hill. He discussed Hill as an outline for evaluating and synthesizing his prior research and the relevant scientific literature that he found while preparing his reports for plaintiffs. Colditz opined that relative risk did not have to be 2.0 for an inference of causation, and provided an example in which the IARC found a relative risk of approximately 1.25 in WHI data about breast cancer a sufficient basis to declare causation. He added that the most fundamental question for any study was how well it was designed to identify the nature and extent of the relevant exposure, and explained that case-control studies that focus on one disease may be superior in that regard to the cohort studies that typically cover too broad a range of diseases or conditions to give them the same attention.

For studies of talc and ovarian cancer, Colditz opined that the most important Hill factors were consistent reports of an association, dose response, and biological plausibility. He assessed the epidemiological studies for bias and confounding, and found that some meta-analyses paid insufficient attention to the "quality" of the measures that their underlying studies used for

talc exposure and for the participants' outcomes. Colditz also considered the results of in vitro experiments.

Defendants' experts stated reasons for considering case-control studies to be unreliable. But the choice of those reasons over those of plaintiffs' experts or of the Manual is a judgment about their relative credibility. For example, while the IARC found only "limited evidence" of an association between perineal talc use and ovarian cancer and expressed general reservations about the limitations of epidemiological studies, it did not find the studies, let alone case-control studies in particular, unsuitable for reliance. Neither the Manual nor Hill requires a study to report a risk or odds ratio of 2.0 to be considered support for an inference of causation. At substantially lower ratios, which they did not quantify, they counseled greater attention to the possibility of bias, confounding, and likely alternative causes.

The cohort, case-control, and pooled or meta-analyses in the record contained considerably more than minimal support for an association of talc with ovarian cancer, whether they are considered together or just by kind of study. The two hospital-based case-control studies (Cramer 1982 and Ness 2000), along with four of the five population-based case-control studies (Chang 1997, Cramer 1999, Mills 2004, and Cramer 2016) and one of the three cohort studies (Gertig 2000), reported a statistically significant association. In

addition, all of the pooled or meta-analyses reported a statistically significant association. While the earlier pooled and meta-analyses called the association weak or doubtful due to variability among the underlying studies (Huncharek 2003) or the lack of a dose response (Langseth 2008), the more recent ones (Terry 2013 and Ness 2015) did not.

The NCI website and some of the studies noted that serous and endometrial ovarian cancer are both subtypes of epithelial ovarian cancer (Cramer 1999, Gertig 2000, Mills 2004, Gates 2009, Terry 2013, Ness 2015). They observed that those and the other subtypes may be different in genesis and behavior, but also that the differences had not yet been established. They named borderline and invasive tumors of each subtype as a separate subtype by itself, they did not contradict the hearing testimony of one defense expert that borderline ovarian tumors "are rarely precursors to" invasive ovarian cancer, and neither Cramer nor Colditz miscited them as if they did.

Among the histological subtypes of epithelial ovarian cancer, four of the studies found the association with talc to be strongest for the serous invasive subtype (Cramer 1999, Gertig 2000, Mills 2004, and Cramer 2016). One of those (Gertig 2000) found a statistically significant association for that subtype only, while noting that studies might have lacked the power to find an association with other subtypes if those cancers have a long latency. Another

one (Cramer 2016) found the association to be strongest between perineal talc use for more than twenty-four years and both the serous invasive and endometrioid subtypes.

The studies provided less support for a dose response. The cohort studies did not state results about it, while one of the two hospital-based case-control studies found a dose response (Ness 2000). Of the population-based case-control studies that found a statistically significant association of talc with ovarian cancer, two found that the dose response was marginal (Chang 1997 and Cramer 1999), one found a dose response for frequency of use but not duration (Cramer 2016), and one found no dose response (Mills 2004). Of the three pooled or meta-analyses that addressed dose response, one found it be minimal (Terry 2013), one found it to be inconsistent (Ness 1999), and one found no "clear" response (Huncharek 2003). Many of the studies noted the inherent difficulty in estimating the amount of product used in any application or of the talc within it (for example, Terry 2013).

Cramer's opinions were substantially consistent with those studies. Cramer applied the Hill factors in discussing the studies on which he relied. He addressed data quality in the meta-analyses, such as the varying classifications of talc use, and he explained that reanalyzing them with a more nearly uniform classification of talc use as meaning only perineal use would

have had little effect on their results. He found the meta-analyses consistent in showing a statistically significant association, including the meta-analysis he performed in preparing his report, which he said had little heterogeneity even though it encompassed both cohort and case-control studies. He acknowledged the limited evidence of a dose response and explained that it could reflect the difficulty of quantifying the amount of talc in each application. For the NCI's conclusion that WHI and NHS did not support an inference of causation, Cramer described what he saw as selection bias in WHI and the failure of NHS to consider changes in the participants' talc use over time.

Colditz opined that the epidemiological studies as a whole showed an increased risk of ovarian cancer from talc use, and that to a lesser degree they supported the inference of a dose response. One pooled analysis with such results was the IARC 2006 report, which in turn relied on two of Cramer's studies and one each from Chang and Ness among what it considered the eight studies of higher quality. For WHI and NHS, Colditz's descriptions of the shortcomings were similar to Cramer's. Colditz also described the extra measures in Terry's 2013 pooled analysis for the OCAC to minimize bias from study heterogeneity.

The FDA found the absence of "conclusive evidence" that talc causes ovarian cancer, based mostly on the lack of general acceptance of a biological

mechanism. However, it did not find the proposed biological mechanisms implausible or contrary to established science, and it called the "growing body" of epidemiological study evidence "difficult to dismiss." One of its reasons for finding the evidence less than conclusive was the possibility that cases of cancer were caused by asbestos in the talc rather than the talc itself. The NCI similarly refrained from calling an association between ovarian cancer and talc or between ovarian cancer and inflammation to be implausible, even though it found the evidence to be inadequate due to inconsistent study results.

Of all the studies, the only ones that reported results for a statistically significant association of inflammation with ovarian cancer were two of the pooled or meta-analyses. One of those found such an association (Trabert and Ness 2014), while the other found it to be inconclusive (Ness 1999).

The only studies with discussions of how talc might cause ovarian cancer in theory were case-control studies. The discussions started with the possibility that migratory talc would cause ovarian inflammation, either directly (Cramer 1982), by causing foreign body entrapment of ovarian surface epithelium (Cramer 1982, Chang 1997, and Cramer 1999), or by getting entrapped in ovulation inclusion cysts (Chang 1997). Two studies discussed later versions of the inflammation hypothesis, which involved the immune

system and reactive oxygen species or mucins (Mills 2004, Cramer and Godleski 2007). Another study, a more recent one that did not focus on talc, discussed how inflammation at sites other than the ovaries could result in ovarian cancer (Trabert 2014).

The record on laboratory testing to connect the presence of talc with ovarian cancer was sparser, but did not contradict it. The presence of talc in ovaries had long been established (Chang 1997, citing published studies from 1961 and 1971; Cramer and Godleski 2007; Langseth 2008). Godleski, whose work and testimony the court named without criticism, found talc in tissue samples of both Carl's and Balderrama's ovaries, but no inflammation. Doctor Lewis Chodosh, an expert for defendants who was a practicing physician, a professor of cancer biology at the University of Pennsylvania School of Medicine, its overseer of faculty research on human carcinogenesis, and an editor of medical journals and member of peer-review panels, agreed that talc can migrate to the ovaries. Omiecinski, whose report and testimony the court likewise refrained from criticizing, explained the possible role of migrating talc in the inflammation hypothesis, and the discussion of that hypothesis in numerous published studies.

Cramer agreed that any causal mechanism must "make sense" in terms of "what is known." He discussed the evidence that talc can migrate to the

ovaries and the development of evidence relating reactive oxygen species and mucins to ovarian cancer. He then explained how it supported his initial view that talc might directly cause changes in ovarian tissue that contribute to carcinogenesis, his later view that talc could contribute to carcinogenesis indirectly by causing inflammation that generates reactive oxygen species, and his current view that talc's contribution could be to chronic inflammation within the upper reproductive tract that eventually blunted the immune system's ability to respond to the markers that an ovarian cancer emits.

Colditz rested his opinion about the biological plausibility of inflammation theories on the work of other experts. Some of those experts established that talc can travel to the ovaries or that talc can cause inflammation, while the epidemiologists who found an association between talc use and ovarian cancer did not see a reason, pending actual demonstrations, why an inflammatory process would be inconsistent with the genesis of ovarian cancer.

On specific causation, Cramer discussed Carl's personal history, her reported talc use, and her alternative known risk factors, primarily obesity and nulliparity. He performed a statistical analysis on a data set that defendants' experts did not challenge, and he found a statistically significant odds ratio of 2.05 for Carl's cancer subtype, serous borderline, among women with as many

perineal applications of talc as Carl. He performed one meta-analysis of studies that considered ovarian cancer in relation to obesity, and another of studies about ovarian cancer and parity, and concluded from their generation of odds ratios lower than 2.05 that talc likely contributed more to Carl's cancer than her obesity or nulliparity did. He then named several other possible risk factors for her and explained how the studies that addressed them failed to show a statistically significant association between them and her tumor subtype.

Cramer performed the same evaluation for Balderrama and her cancer subtype, endometrioid. That included meta-analyses of the studies of perineal talc use and of her other known risk factors with endometrial ovarian cancer. He found a statistically significant odds ratio of 1.79 for her cancer subtype among women with at least approximately sixty percent as many perineal applications of talc as Balderrama reported, and he found that to be higher than the ratio for her other main risk factors. Cramer acknowledged that was not the case for Balderrama's uterine endothelial tumor, and he explained why it was a separate primary cancer rather than an incident of her ovarian cancer. Cramer's findings for Carl's and Balderrama's subtypes of ovarian cancer were consistent with the results in his 2016 case-control study.

Colditz did not opine on specific causation, but he noted that four of the five studies in the IARC 2006 report that addressed subtypes found the risk increase to be greatest for serous ovarian cancer. He added that Gertig in 2000 found that stratification of the NHS data by subtype showed a significant increase in risk for serous invasive cancer, and that Gates in 2008 found the risk for invasive serous cancer to be somewhat stronger than for ovarian cancer without regard to subtype.

VI.

The trial judge was called upon to assess whether the opinions were the product of reliable data and employed methodologies accepted by the scientific community. Instead, he selected defendants' scientific methodologies over plaintiffs', a process well beyond the gatekeeping function, and which resulted in an abuse of discretion. Under prior law or post-Accutane, the court erred by categorically characterizing cohort studies as more credible than case-control studies; imposing a relative risk of 2.0 as the threshold for the result of an epidemiological study to become reliable for any purpose; requiring Cramer and Colditz to develop their own studies to support their inflammation hypotheses instead of relying on the work of other experts; and requiring Cramer and Colditz to disprove the causation theories of defendants' experts.

Furthermore, the trial judge, as to specific causation, erred by mischaracterizing Cramer's methodology, which was unobjectionable.

The judge also erred because he described the Manual, incorrectly, as characterizing case-control studies generally as subject to informational bias. Nor did the Manual admonish users about the superiority of studies with large samples. Nothing in the Manual imposed a threshold for a sample size to be "large enough"; in fact, all the case-control studies in the record had sample sizes in the hundreds or thousands. The judge did not identify errors that would make it unsound for an expert to rely on these studies that the relevant scientific field accepted for that purpose.

The case-control studies were a substantial portion of the hearing record, and defendants' experts did not suggest that they were an insubstantial portion of the entire relevant scientific record. The case-control studies here consistently reported statistically significant associations of talc with ovarian cancer, as did one of the three cohort studies and the two most recent of the five pooled or meta-analyses. Some of the pooled or meta-analyses included both cohort and case-control studies, and they did not report a need to adjust for perceived inferiorities of the latter. Furthermore, the five studies in this record that were among the eight on which the IARC focused in its 2006 report, due to their "higher quality," were all case-control studies.

Cramer's use of statistical analyses for each plaintiff's cancer, to account for the contribution of talc, was consistent with the methodologies of the numerous published studies in the record. Defendants' experts conceded the migration of talc to the ovaries, and studies on which the judge himself relied provided evidence of an inflammatory effect. The judge's suspicions regarding Cramer's conclusions were therefore a judgment regarding their credibility.

The judge contrasted the willingness of plaintiffs' experts to testify in 2016 that the legal standard had been satisfied with their prior reluctance to conclude that the evidence of talc's association with ovarian cancer constituted scientific proof. Accordingly, he opined that Cramer relied on a "made-for-litigation methodology" and Colditz issued an "ipse dixit[.]" But the legal standard that governed the Rule 104 hearing and decision is not absolute scientific proof. The issue is methodology, and the reliability of the data upon which the work relied.

Defendants' experts generally challenged plaintiffs' experts' inflammation hypotheses, offering alternative biological mechanisms for ovarian cancer that did not involve talc. It is not improper for a court to expect an expert to demonstrate the soundness of his or her methodology "from the perspective of others within the relevant scientific community." Accutane, 234 N.J. at 399-400. When "the relevant scientific literature contains evidence

tending to refute the expert's theory," the expert may not decline to "acknowledge or account for" it. Rezulin, 369 F. Supp. 2d at 425 (elaboration of the point cited by Accutane, 234 N.J. at 400).

The judge adopted evidence from defendants' experts about talc's ameliorative effect on lung cancer as if it had been proven generally for all solid cancers including ovarian cancer. However, no laboratory research in the hearing record demonstrated that lung and ovarian cancer are similar, particularly in their responses to talc, and all the experts agreed that a carcinogen could cause cancer in some organs but not others.

Cramer's report relied on a laboratory research study regarding the inflammatory effect on ovarian cells when talc is placed directly upon them. The judge ignored that finding despite attaching a summary of that study to his opinion. In addition, the judge relied on the absence of an association between talc and other cancers of the reproductive tract to conclude that the inflammation hypothesis was invalid, when the record did not establish that the association's absence and the hypothesis were irreconcilable.

The judge accepted the defense experts' opinion that mutations in critical genes is the mechanism that causes cancer, and hence since talc does not cause mutations, it cannot cause cancer. Although a factfinder can certainly accept all, some, or none of an expert's findings, City of Long Branch v. Liu, 203 N.J.

464, 491 (2010), that was not the judge's role at the Rule 104 hearing. His task was to assess the soundness of the methodology of plaintiff's experts and the soundness of the "underlying data and information." Accutane, 234 N.J. at 390. Instead, he chose between plaintiffs' and defendants' experts based on his assessment of the credibility of their opinions.

We are satisfied that plaintiffs' experts adhered to methodologies generally followed by experts in the field, and relied upon studies and information generally considered an acceptable basis for inclusion in the formulation of expert opinions. Suppression of their testimony was an abuse of discretion.

That reversal means there is a dispute of material fact. Thus, summary judgment dismissing plaintiffs' complaints must also be reversed. See R. 4:46-2(c).

Reversed. We do not retain jurisdiction.