

NOT FOR PUBLICATION WITHOUT THE APPROVAL
OF THE APPELLATE DIVISION

SUPERIOR COURT OF NEW JERSEY
APPELLATE DIVISION
DOCKET NO. A-4698-14T1
A-0910-16T1

IN RE: ACCUTANE LITIGATION

APPROVED FOR PUBLICATION

July 28, 2017

APPELLATE DIVISION

Argued March 7, 2017 – Decided July 28, 2017

Before Judges Reisner, Koblitz and Sumners.

On appeal from the Superior Court of New Jersey, Law Division, Atlantic County, Case No. 271 (MCL).

Bruce D. Greenberg and David R. Buchanan argued the cause for appellants (Seeger Weiss, LLP, Lite, DePalma, Greenberg, LLC, and Weitz & Luxenberg, PC, attorneys; Mr. Buchanan and Peter Samberg, of counsel; Mr. Buchanan, on the briefs).

Paul W. Schmidt (Covington & Burling, LLP) of the District of Columbia bar, admitted pro hac vice, argued the cause for respondents Hoffman LaRoche, Inc. and Roche Laboratories, Inc. (Gibbons PC, Dughi Hewit & Domalewski, PC and Mr. Schmidt, attorneys; Michelle M. Bufano, Natalie H. Mantell, Russell L. Hewit, Mr. Schmidt and Michael X. Imbroscio (Covington & Burling, LLP) of the District of Columbia bar, admitted pro hac vice, of counsel; Ms. Bufano, on the brief).

Hollingsworth LLP, attorneys for amicus curiae Pharmaceutical Research and Manufacturers of America (Gregory S. Chernack, of counsel and on the brief).

The parties have not filed briefs in A-0910-16.

The opinion of the court was delivered by
REISNER, P.J.A.D.

Plaintiffs, in these 2076 multicounty litigation (MCL) products liability cases, alleged that they developed Crohn's disease¹ as a result of taking Accutane (isotretinoin), a prescription acne drug manufactured by defendants Hoffman-La Roche Inc. and Roche Laboratories Inc. (collectively Roche or defendants). After a Kemp² hearing, the trial court issued a February 20, 2015 order granting defendants' omnibus motion to bar plaintiffs' experts - Dr. David Madigan, a statistician, and Dr. Arthur Asher Kornbluth, a gastroenterologist - from testifying, among other things, that the epidemiology studies on which the defense relied were flawed and unreliable, and that Accutane can cause Crohn's disease. The trial court also directed the parties to prepare an order listing the lawsuits affected by the ruling, and subsequently issued a May 8, 2015 order dismissing 2076 MCL

¹ Crohn's disease is a form of inflammatory bowel disease (IBD).

² Kemp ex rel. Wright v. State, 174 N.J. 412, 417 (2002).

claims with prejudice. Plaintiffs appeal from those orders.³

On this appeal, plaintiffs primarily contend that the trial court misapplied its discretion in finding that the methodologies Madigan and Kornbluth used were scientifically unreliable and inadmissible. After reviewing the record, we reverse the orders on appeal and remand this case to the trial court.

We agree with plaintiffs that the trial court went beyond its gatekeeping function, as set forth in Rubanick v. Witco Chemical Corp., 125 N.J. 421, 449 (1991), Landrigan v. Celotex Corp., 127 N.J. 404 (1992), and Kemp, supra, 174 N.J. at 412.⁴ The trial court took too narrow a view in determining whether the experts were using accepted scientific methodologies to analyze the evidence, and improperly determined the weight and credibility of the experts' testimony. Among other things, the judge

³ In a second appeal (A-0910-16), ninety-eight plaintiffs appeal from a September 19, 2016 order dismissing their complaints on the same basis. By order dated December 7, 2016, we granted an unopposed motion to consolidate A-0910-16 with the current appeal, A-4698-14 (the first appeal); however, we excused the parties in the second appeal from filing briefs or appendices, based on their agreement to be bound by the outcome of the first appeal.

⁴ Plaintiffs also argue that the trial court erred as a matter of law in applying the strict, scientific certainty admissibility standard, instead of the relaxed standard set forth in Rubanick. That argument is without sufficient merit to warrant discussion. R. 2:11-3(e)(1)(E).

inappropriately condemned the experts for relying on relevant scientific evidence other than epidemiological studies, despite their plausible explanations for doing do.⁵ Consequently, we conclude that the trial court mistakenly exercised discretion in barring the experts' testimony.

In reaching our conclusion, we emphasize that we are not placing this court's imprimatur on plaintiffs' experts or on their opinions. The experts on both sides are highly reputable scientists, who view the evidence differently. We find no basis to describe plaintiffs' experts pejoratively as "hired guns," any more than the defense experts are "hired guns." Their testimony should not have been barred because their analyses emphasized different evidence and produced different conclusions than those reached by the defense experts. The fact that plaintiffs' experts found certain evidence to be critically important did not constitute improper "cherry picking," because they provided plausible scientific explanations for their choices. See State v. Dreher, 302 N.J. Super. 408, 464 (App. Div. 1997) ("Expert testimony should not be excluded merely because it fails to account

⁵ Those same types of evidence were held admissible by a prior judge, who had handled the Accutane MCL litigation for a decade.

for some condition or fact that the opposing party considers relevant.").

We are not predicting whether a jury will find plaintiffs' experts - or defendants' experts - credible or persuasive. That is not our role, as it was not the trial court's role in the Kemp hearing. See Hisenaj v. Kuehner, 194 N.J. 6, 24 (2008) (N.J.R.E. 104 hearings "are intended to determine admissibility, not credibility."). We only hold that, on the record created in the Kemp hearing in this case, the plaintiffs' experts provided well-explained scientific reasons for analyzing the available evidence differently from the defense experts, and for relying more heavily on different evidence than the defense experts relied on.

Accordingly, plaintiffs are entitled to present the experts' testimony at trial.

I

This case cannot be viewed in a vacuum. It is one in a long series of mass tort litigations concerning Accutane.⁶ We need not

⁶ McCarrell v. Hoffman-La Roche, Inc. (McCarrell I), No. A-3280-07 (App. Div. Mar. 12, 2009), certif. denied, 199 N.J. 518 (2009); Kendall v. Hoffman-La Roche, Inc. (Kendall I), No. A-2633-08 (App. Div. Aug. 5, 2010), aff'd, 209 N.J. 173 (2012); Sager v. Hoffman-La Roche, Inc., No. A-3427-09 (App. Div. Aug. 7, 2012), certif. denied, 213 N.J. 568 (2013); Gaghan v. Hoffman-La Roche, Nos. A-2717-11, A-3211-11, A-3217-11 (App. Div. Aug. 4, 2014); McCarrell

review the history in detail, as it is set forth in a series of previous unpublished opinions issued by different panels of this court. We summarize only what is important to this case.

For more than a decade, the same trial judge had handled the Accutane cases. To some extent, that judge's familiarity with the prior litigation, and with the multiplicity of scientific issues involved, may have shaped the way the parties and their experts prepared for the current litigation.⁷ The first judge's rulings no doubt also shaped the parties' litigation strategies.

In particular, during the course of the litigation, the first judge determined that the opinions of plaintiffs' experts, based on the same types of evidence relied on by plaintiff's experts in this case, would be admissible as scientifically reliable. We

v. Hoffman-La Roche, Inc. (McCarrell II), No. A-4481-12 (Aug. 11, 2015), rev'd and remanded, 227 N.J. 569 (2017); Kendall v. Hoffman-La Roche, Inc. (Kendall II), No. A-0301-14 (June 16, 2016); and Rossitto v. Hoffman-La Roche, Nos. A-1236-13, A-1237-13 (July 22, 2016), certif. denied, 228 N.J. 419 (2016).

⁷ Both of the parties' epidemiology experts (Dr. Madigan and Dr. Steven N. Goodman) who had testified previously, expressed their belief that, to some extent, their current reports and testimony would be viewed in light of their testimony in previous Accutane trials.

affirmed that determination in McCarrell I, supra, A-3280-07,⁸ finding that animal studies, case reports, analogous medications, and other evidence relied on by plaintiffs' experts, were types of evidence accepted in the scientific community.⁹

In the present case, defendants contend that the existence of epidemiological studies now precludes reliance on the other types of evidence on which plaintiff's experts had previously relied. However, the studies on which defendants rely are not the controlled clinical trials that the Federal Judicial Center's Research Manual on Scientific Evidence calls "the gold standard" of scientific evidence. Rather they are observational studies that

⁸ Unpublished opinions are not to be cited as legal precedent, and we do not do so here. R. 1:36-3. However, it is appropriate to consider an unpublished opinion of this court where, as here, it forms a part of the history of the case on appeal. See Mountain Hill, L.L.C. v. Twp. Comm. of Twp. of Middletown, 403 N.J. Super. 146, 155 n.3 (App. Div. 2008), certif. denied, 199 N.J. 129 (2009). Moreover, an unpublished opinion of this court is binding on the trial court in the same case. Ibid. The parties have not briefed and, hence, we do not decide, whether an unpublished opinion of this court is binding on the trial court in the same MCL docket, albeit in a different case within that docket.

⁹ In a 2014 oral opinion, the first trial judge made a detailed analysis of similar testimony by Dr. David Sachar and Dr. Madigan concerning the connection between Accutane and ulcerative colitis. The first judge concluded that the expert testimony, which relied on very similar types of evidence as that used in this case, was admissible. That decision was appealed, but was settled before we decided the appeal. See Kendall II, supra, A-0301-14.

depend on the collection of information from databases or from patient questionnaires. Plaintiffs' experts testified that the studies are biased and otherwise flawed. We conclude that plaintiffs should be entitled to present that testimony at trial, along with their affirmative evidence in support of their case.

II

We begin with some background as to Accutane, the epidemiological studies of the drug, and relevant scientific principles of epidemiology.

A. Accutane

In 1982, the Food and Drug Administration (FDA) approved defendants' application to market Accutane, the brand name for isotretinoin, "to treat recalcitrant nodular acne that has not responded to other regimens." Kendall I, supra, 209 N.J. at 180. The drug is a retinoid, derived from vitamin A, and is very effective in treating severe acne. Ibid. It is well established that Accutane "has a number of known side effects, including dry lips, skin and eyes; conjunctivitis; decreased night vision; muscle and joint aches; elevated triglycerides; and a high risk of birth defects if a woman ingests the drug while pregnant." Ibid. There is also some evidence that Accutane, which was

originally studied for use in treating cancer, has an effect on the gastrointestinal tract.

The MCL cases concern the alleged propensity of Accutane to cause IBD, a chronic disease which primarily manifests as one of two diseases: Crohn's disease or ulcerative colitis. Id. at 180-81. Although both ulcerative colitis and Crohn's disease share the same core symptoms, including abdominal pain, frequent and often bloody bowel movements, and rectal bleeding, there are differences in the clinical presentation of the disease and the triggers statistically associated for developing it, which include family history, infections, frequent use of some antibiotics, smoking, and possibly the use of oral contraceptives and nonsteroidal anti-inflammatory drugs. Id. at 181.

The peak onset of the disease occurs during adolescence—the same period that individuals are likely to have been prescribed Accutane. Ibid. For both diseases there may be a significant latency effect (the time from the exposure to the trigger for IBD to the first symptom of the disease) and a prodromal period (the time from the first symptom of the disease to diagnosis).

B. Epidemiological studies

For the first six years of this MCL litigation, from 2003 to

2009, there were no epidemiological studies regarding Accutane and IBD. In previous trials, the plaintiffs were permitted to present expert testimony that relied on animal studies, human clinical studies, case reports, class effects, published scientific literature, causality assessments, and biological plausibility. McCarrell I, supra, slip op. at 86; Kendall I, supra, slip op. at 85-86; Sager, supra, slip op. at 20.

The first two epidemiological studies (Crockett and Bernstein),¹⁰ were published in 2009 and in 2010, finding no statistically significant increased risk of developing Crohn's disease from the use of Accutane, although the Crockett study found ulcerative colitis is associated with exposure to the drug. The Crockett and Bernstein studies were addressed in expert testimony in Gaghan, McCarrell II, and Rossitto. In Kendall II, the expert witnesses addressed four new epidemiological studies

¹⁰ Seth D. Crockett et al., Isotretinoin Use and the Risk of Inflammatory Bowel Disease: A Case-Control Study, 105 Am. J. Gastroenterol. 1986 (Sept. 2010) (ulcerative colitis but not Crohn's disease is associated with isotretinoin use); Charles N. Bernstein et al., Isotretinoin is not Associated with Inflammatory Bowel Disease: A Population-Based Case-Control Study, 104 Am. J. Gastroenterol. 2774 (Nov. 2009) (unlikely that isotretinoin use is associated with development of IBD).

(Etminan, Alhusayen, Fenerty, and Racine).¹¹ After the trial in Kendall II, two additional studies were published (Rashtak and Sivaraman).¹² The epidemiological studies vary in whether they show that Accutane increases or decreases the risk of developing Crohn's disease. However, with one exception, none of them demonstrates a statistically significant increased risk of developing Crohn's disease from exposure to Accutane. One small study (Sivaraman) did find a statistically significant increased risk. However, when the study authors adjusted the study results for antibiotic use, the results were no longer statistically significant. Plaintiffs' experts questioned the appropriateness of that adjustment.

C. Epidemiology

¹¹ Mahyar Etminan et al., Isotretinoin and Risk for Inflammatory Bowel Disease, 149 JAMA Dermatol. 216 (Feb. 2013) (no increased risk for IBD); Raed O. Alhusayen et al., Isotretinoin Use and the Risk of Inflammatory Bowel Disease: A Population-Based Cohort Study, 133 J. Invest. Dermatol. 907 (2013); Sarah Fenerty et al., Impact of Acne Treatment on Inflammatory Bowel Disease, 68 J. Am. Acad. Dermatol. 6751 (Apr. 2013); Antoine Racine et al., Isotretinoin and Risk of Inflammatory Bowel Disease: A French Nationwide Study, 109 Am. J. Gastroenterol. 563 (Apr. 2014).

¹² Shadi Rashtak et al., Isotretinoin Exposure and Risk of Inflammatory Bowel Disease, 150 JAMA Dermatol. 1322 (Dec. 2014); Susil Silverman et al., Risk of Inflammatory Bowel Disease from Isotretinoin: A Case Control Study (Oct. 2014).

In understanding the epidemiological studies, it is first helpful to define the methodology used in conducting such studies and the relevant terms, as testified by the experts at the hearing and as set forth in the Federal Judicial Center, Reference Manual on Scientific Evidence 549, 555 (3d. ed. 2011) (Reference Manual or Manual).¹³ "Epidemiology is the field of public health and medicine that studies the incidence, distribution, and etiology of disease in human populations." Id. at 551. "Epidemiology assumes that disease is not distributed randomly in a group of individuals and the identifiable subgroups, including those exposed to certain agents [such as prescription drugs], are at increased risk of contracting particular diseases." Ibid. Epidemiological studies identify agents that are associated with an increased risk of a disease in groups of individuals, but "is not equivalent to causation." Id. at 552.

There are two types of epidemiological studies: experimental and observational. Id. at 555. Experimental studies, or double-blind randomized control trials, in which one group is exposed to

¹³ Available at <https://www.fjc.gov/sites/default/files/2015/SciMan3D01.pdf>. The epidemiological section was written by Michael D. Green, among others, and is entitled Reference Guide on Epidemiology.

an agent and the other is not, are "considered the gold standard for determining the relationship of an agent to a health outcome or adverse side effect." Ibid. There are, however, no Accutane experimental studies because although such studies have the potential to provide higher quality evidence, they cannot ethically be conducted if researchers suspect that a drug's side-effects are harmful. Id. at 555-56.

Instead, all of the Accutane epidemiological studies to date are less rigorous observational studies, which are considered to be the next best available evidence. Id. at 556. There are two types of observational studies: 1) a case-control study, which measures and compares the frequency of exposure in the group with the disease (cases) and a similar group without the disease (controls); and 2) a cohort study, which compares a group of exposed and unexposed individuals over a period of time. Id. at 557-59. In these studies, researchers "observe" individuals who have already been exposed to the drug and compare them to a group of individuals who have not. Id. at 555-56.

Unlike experimental studies in which risk factors can be controlled, observational studies generally focus on individuals living in a community, "for whom characteristics other than the

one of interest, such as diet, exercise, exposure to other environmental agents, and genetic background, may distort a study's results." Id. at 556. "[T]he Achilles' heel of observational studies is the possibility of differences in the two populations being studied with regard to risk factors other than exposure to the agent." Ibid.

Epidemiological studies commonly express the strength of association between exposure to a drug and a disease in numerical terms as: 1) "relative risk" (RR), the ratio of the incidence rate of a disease in exposed individuals to the risk among the unexposed; or 2) "odds ratio" (OR), the ratio of the odds that an individual with the disease was exposed to the drug to the odds that an individual without the disease was exposed. Id. at 566-69.¹⁴ An RR of 1.0 means that the relative risk is equal to the "null hypothesis," that is, that the risk in individuals exposed to Accutane is the same as the risk in unexposed individuals, or that Accutane use is not associated with an increased risk of

¹⁴ The Manual explains that an OR is "a convenient way to estimate the relative risk in a case-control study when the disease under investigation is rare." Id. at 568. Most of the studies at issue in this case are case-control studies. An OR is calculated somewhat differently in a cohort study but the difference is not pertinent here.

developing Crohn's disease. Id. at 567. If the RR is greater than 1.0, the risk in exposed individuals is greater than the risk in unexposed individuals. Ibid. For example, an RR of 1.5 means that an exposed individual has a 50% greater chance of contracting Crohn's disease. If the RR is less than 1.0, the exposed group has a decreased risk of contracting the disease. Ibid. Thus, an RR of .32 represents a 68% reduction in risk, which might mean that the drug had a protective effect on developing the disease.

The OR or RR is, however, only an estimate of the true value. Determining whether an association identified in an epidemiological study is causal "requires an understanding of the strengths and weaknesses of the study's design and implementation, as well as a judgment about how the study findings fit with other scientific knowledge." Id. at 553. An assessment must be made of the power of the study to detect associations, the role of chance, and what sources of error might have produced a false result, including sampling variability, bias, and confounding variables (extraneous variables that may affect result). Id. at 566-97.

Therefore, a showing of an increased relative risk for Crohn's disease does not automatically prove that Accutane use creates a

higher risk of developing the disease because the discrepancy between the exposed and unexposed groups could be the product of chance as a result of random sampling error. Id. at 553. In determining whether a relative risk greater than 1.0 is a true association or the result of random error, researchers consider whether the association is statistically significant. Id. at 628. In making that assessment, researchers calculate a p-value, which "represents the probability that an observed positive association could result from random error even if no association were in fact present." Id. at 576. The p-value quantifies the statistical significance of a relationship; the smaller the p-value the greater the likelihood that associations determined in a study do not result from chance. Id. at 626. The most commonly used p-value is .05, that is for example, that there is a 5% chance that the relative risk could have occurred by random error. Id. at 576-77.

A more sophisticated approach, which was used in the studies at issue in this case, involves calculating a confidence interval (CI):

A confidence interval is a range of possible values calculated from the results of a study. If a 95% confidence interval is specified, the range encompasses the results we would expect

95% of the time if samples for new studies were repeatedly drawn from the same population. . . . The advantage of a confidence interval is that it displays more information than significance testing. "Statistically significant" does not convey the magnitude of the association found in the study or indicate how statistically stable that association is. A confidence interval shows the boundaries of the relative risk based on selected levels of . . . statistical significance. . . . [T]he confidence interval reveals the likely range of risk estimates consistent with random error.

[Id. at 580.]

If, for example, a study reveals a RR of 1.5, which represents an elevated risk of developing Crohn's disease, that result might or might not be considered statistically significant, depending on the boundaries of the confidence interval. If the CI includes 1.0 (the null hypothesis, meaning that taking Accutane neither increases nor decreases the risk of developing Crohn's disease), then the 1.5 result is said not to be statistically significant. However, if the CI is entirely above 1.0, for example if it ranges from 1.2 to 3.2, then the 1.5 RR would be considered statistically significant. Id. at 580-81.

In assessing whether the failure of a study to find a statistically significant association was exonerative of the drug or simply inconclusive, scientists consider the "power" of a study,

or "the probability of finding a statistically significant association of a given magnitude (if it exists) in light of the sample sizes used in the study." Id. at 582. "The power of a study depends on several factors: the sample size; the level of statistical significance specified; the background incidence of disease; and the specified relative risk that the researcher would like to detect." Ibid. The higher the power of the study the less likely it will show a false negative. Ibid. For example, a study with a likelihood of .25 of failing to detect a true RR of 2.0, has a power of .75, meaning the study has a 75% chance of detecting a true RR of 2.0. Ibid. On the other hand, a study with low power has a greater likelihood of failing to detect a significant relative risk, even though such a risk exists. "With large numbers [of individuals included in the study group], the outcome of the test is less likely to be influenced by random error, and the researcher would have greater confidence in the inferences drawn from the data." Id. at 576.

Under the proper circumstances, researchers can increase the power of a series of studies by conducting a meta-analysis, which involves pooling the results of different studies, some of which are small and lack statistical power, to arrive at a single figure

to represent the totality of the studies. Id. at 608. The Manual indicates, however, that a meta-analysis may produce an unreliable result.

The appeal of a meta-analysis is that it generates a single estimate of risk (along with an associated confidence interval), but this strength can also be a weakness, and may lead to a false sense of security regarding the certainty of the estimate. A key issue is the matter of heterogeneity of results among the studies being summarized. If there is more variance among study results than one would expect by chance, this creates further uncertainty about the summary measure from the meta-analysis. Such differences can arise from variations in study quality, or in study populations or in study designs. Such differences in results make it harder to trust a single estimate of effect; the reasons for such differences need at least to be acknowledged and, if possible, explained. People often tend to have an inordinate belief in the validity of the findings when a single number is attached to them, and many of the difficulties that may arise in conducting a meta-analysis, especially of observational studies such as epidemiologic ones, may consequently be overlooked.

[Ibid.]

III

We next address the parties' conflicting testimony on the subjects of gastroenterology and epidemiology. As background, the

following chart¹⁵ summarizes the epidemiological studies at issue in this case:

STUDY	DATABASE AND SUBJECTS	NO. OF SUBJECTS	RR for CD at CI 95%	ACCUTANE EXPOSURE PRIOR TO DIAGNOSIS	STUDY RESULTS
Bernstein 2009 manuscript (case-control)	Canadian Health Ins.	21,500 (total) 1118 (CD)	1.15 (0.61-2.02)	2.6 years	Positive association (increased risk), but not SS
Crockett 2010 manuscript (case-control)	US Health Ins. (55 million)	29,000 (total) 3664 (CD)	0.68 (0.28-1.68)	1 year	Negative association (decreased risk)
			0.89 (0.32-2.52)	2 year	Negative association (decreased risk)
Etminan 2013 manuscript (case-control) (meta-analysis)	US Health Ins. (women who had taken oral contraceptives)	45,000 (total) 1103 (CD)	1.05 (0.5501.98)	1 year	Positive unadjusted association
			0.91 (0.37-2.25)		Negative adjusted association (decreased risk)
			0.75 (0.46-1.24)		Negative association (meta-analysis)
Racine 2014 manuscript (case-control)	French Health Ins. (47 million)	44,000 (total) 2829 (CD)	0.45 (0.24-0.85)	1 to 2 years	SS protective association (reduced risk)
Alhusayen 2013 manuscript (cohort)	Canadian Database (4.5 million)	46,922 (treated with Accutane)	1.17 (0.90-1.52)	1 year	Positive association (increased risk), but not SS

¹⁵ For purposes of the chart, we abbreviate Crohn's disease as "CD" and statistical significance as "SS."

STUDY	DATABASE AND SUBJECTS	NO. OF SUBJECTS	RR for CD at CI 95%	ACCUTANE EXPOSURE PRIOR TO DIAGNOSIS	STUDY RESULTS
Sivaraman 2014 abstract/ poster (case-control)	US Patient Questionnaire from three clinics	509 (total)	5.6 (1.1-28.0) 4.8 (0.3-70)	unknown	Positive unadjusted association and SS for CD Positive adjusted association, but not SS for CD
Fenerty 2013 abstract/ power-point (case-control)	Marketscan Medicaid Database	176,889 (total) 324 (CD)	For IBD 0.57 (0.28-1.16) Not reported for CD		Negative association for IBD (decreased risk)
Rashtak 2014 manuscript (cohort)	Mayo Clinic patients	1078 (total)	For IBD 0.28 (0.10-0.79) Not reported for CD		Negative association for IBD (decreased risk)

A. Gastroenterology experts

1. Dr. Arthur Asher Kornbluth

Plaintiffs' expert, Dr. Kornbluth, was a highly qualified expert who was board-certified in internal medicine and gastroenterology and was a professor of medicine at Mount Sinai, the preeminent hospital for IBD. He had specialized in Crohn's disease for twenty-seven years, conducted research on IBD, conducted clinical trials on several drugs intended for use in the management of IBD, treated between 5,000 and 10,000 patients with Crohn's disease, been retained as a consultant to pharmaceutical

companies, and published more than 100 articles on IBD in peer-reviewed scientific journals, textbooks, and other publications.

Kornbluth opined that Accutane can cause Crohn's disease in humans. In reaching that conclusion, Kornbluth relied on his personal experience in treating thousands of patients with the disease. Additionally, he relied on some of the same evidence that Dr. David Sachar, a previous plaintiffs' expert, had relied on in seven previous Accutane trials in this MCL docket,¹⁶ including animal studies, case reports, class effects of Vesanoid, biological plausibility, scientific articles, internal studies, causality assessments, and epidemiological studies. However, as more fully discussed infra, both Dr. Kornbluth and Dr. Madigan testified that most of the epidemiological studies done to date were fundamentally flawed, thus warranting greater reliance on other forms of scientific evidence.

Evidence of an Association

In accord with what he testified was the established scientific methodology, Kornbluth first considered whether there was an association between Accutane and Crohn's disease. In making

¹⁶ McCarrell I and II, Kendall I and II, Sager, Gaghan, and Rossitto.

that determination, he reviewed scientific articles, MedWatch reports, epidemiological studies, and causality assessments, which he found reflected a strong association between Accutane and Crohn's disease.

a. Scientific articles

Kornbluth reviewed articles published in peer-reviewed scientific journals, many of which analyzed a single anecdotal case report, which he found supported a finding that there was an association between Accutane and Crohn's disease.¹⁷ For example, an article by Reddy and several colleagues reported that of the

¹⁷ P. Martin et al., Isotretinoin-Associated Proctosigmoiditis, 93 Gastroenterology 606 (1987) (case report); Philippe Deplaix et al., Acute Hemorrhagic Colitis Probably Due to Isotretinoin with Recurrence Following Reintroduction of Treatment, 20 Gastroenterol. Clin. Biol. 113 (1996) (case report); Marianne Melki et al., Granulomatous Colitis Probably Due to Isotretinoin, 25 Gastroenterol. Clin. Biol. 433 (2001) (case report); J.L.M. Passier et al., Isotretinoin-Induced Inflammatory Bowel Disease, 64 Neth. J. Med. 52 (Feb. 2006) (case reports); Deepa Reddy, M.D. et al., Possible Association Between Isotretinoin and Inflammatory Bowel Disease, 101 Am. J. Gastroenterol. 1569 (July 2006) (adverse event reports); Cristiano Spada, M.D. et al., Isotretinoin-associated Pan-enteritis, 42 J. Clin. Gastroenterol. 923 (Sept. 2008) (case report); Matthew Shale et al., Isotretinoin and Intestinal Inflammation: What Gastroenterologists Need To Know, 58 Gut 737, 739 (June 2009) (study of adverse event reports concluding isotretinoin may act as trigger for IBD in susceptible patients); and Michael B. Brodin, M.D., Inflammatory Bowel Disease and Isotretinoin, 14 J. Am. Acad. Dermatol. 843 (1986) (letter to editor).

approximately four or five million people that took Accutane between 1997 and 2002, the FDA received eighty-five reports of IBD. Using the Naranjo ADR probability scale, the authors found that "4 cases (5%) scored in the 'highly probable' range for isotretinoin as the cause of IBD, 58 cases (68%) were 'probable,' 23 cases (27%) were 'possible,' and no cases were doubtful." Reddy, supra note 17, at 1571. The authors concluded that "isotretinoin appears to be a potential precipitant of IBD." Id. at 1572.

b. MedWatch reports

Kornbluth next reviewed a series of MedWatch reports, reports which are made by physicians, patients and others to the FDA listing among other information, a description of the adverse event and whether it abated after the patient stopped using Accutane and returned after reintroduction (referred to as challenge/dechallenge/rechallenge). He found that if corrected for underreporting, the number of MedWatch reports suggested a strong association between Accutane and Crohn's disease.

c. Epidemiological studies

Kornbluth also reviewed six observational epidemiological studies (Bernstein, Crockett, Alhusayen, Etminan, Racine, and

Sivaraman), only one of which (Sivaraman) found a statistically significant positive association (before adjustment for antibiotic use) between Accutane and Crohn's disease, one found a statistically significant negative association (Racine) and no study concluded that Accutane use presents an increased risk for developing Crohn's disease.¹⁸ Kornbluth opined that despite "significant flaws" in five of those studies that distorted the results, the studies nonetheless provided some evidence of an association between Accutane and Crohn's disease.

Kornbluth relied on the unadjusted results of the Sivaraman study, which was summarized in a published abstract (not manuscript form), and selected by the American College of Gastroenterology to be presented as a poster at their annual conference to enable peers and colleagues to discuss the findings with the researchers. In that small study (509 patients), the authors initially found a statistically significant association between Accutane and Crohn's disease: that the risk of developing the disease was more than five times higher in the group exposed to Accutane. The authors

¹⁸ The studies yielded different results for ulcerative colitis: Crockett found a statistically significant association; Bernstein, Etminan, Racine, and Alhusayen found a positive association, but not a statistically significant association; and Fenerty and Rashtak found a negative association.

collected data using a questionnaire, which included information about use of antibiotics, family history of IBD, and smoking, thereby accounting for confounding variables. The authors then "adjusted" the analysis to remove subjects who had taken antibiotics, which Kornbluth said still yielded a "very striking increased risk" of developing Crohn's disease from Accutane use, although the adjusted sample size was too small to demonstrate statistical significance. The authors concluded that "isotretinoin exposure does not appear to confer risk for Crohn's disease independent of antibiotic exposure." Kornbluth and Madigan both questioned the basis for the adjustment the authors made.¹⁹

Kornbluth opined that the results of the other five studies (Bernstein, Crockett, Alhusayen, Etminan, and Racine), were inconclusive because they: failed to account for the prodrome (see section 1 below); were insufficiently "powered"; or contained design flaws that biased or distorted the results to show a reduced risk of developing Crohn's disease. Nonetheless, Kornbluth found

¹⁹ The authors noted that taking antibiotics alone did not appear to affect the risk of developing Crohn's disease. Nonetheless, they removed the subjects who had taken antibiotics when they recalculated the study results.

that the studies were informative in determining causation, noting that four of the studies (Bernstein, Etminan (unadjusted), Alhusayen, and Sivaraman) found a positive association between the drug and the disease.

1. Prodrome

Kornbluth opined that four of the studies (Crockett, Alhusayen, Etminan, and Racine) had not followed patients for long enough to detect an effect from Accutane exposure and thus had failed to account for the prodrome of Crohn's disease, that is, the delay between the time of the first or early symptoms and the diagnosis. He opined, based on his decades of experience as a treating gastroenterologist, that the average prodrome for Crohn's disease was from two to four years. He found support for that opinion in several studies, including: 1) the Pimentel study,²⁰ a referral-based study (45 of the 66 subjects had Crohn's disease), in which the authors found the mean prodrome for Crohn's disease was 6.9 years; and 2) the Barratt study,²¹ in which the authors

²⁰ Mark Pimentel, M.D. et al., Identification of a Prodromal Period in Crohn's Disease but Not Ulcerative Colitis, 95 Am. J. Gastroenterol. 3458 (Dec. 2000).

²¹ S.M. Barratt et al., Prodromal Irritable Bowel Syndrome May Be Responsible For Delays In Diagnosis In Patients Presenting With

found the mean prodrome was four years. He distinguished the findings of a larger study by Chouraki²² in which the authors found a three-month prodrome, based upon the selection of the patients for the study and the use of patient charts as opposed to patient questionnaires, as used in the Barratt and Pimentel studies.

Kornbluth explained that because the prodrome for Crohn's disease is two to four years, a study that looks back only one-year from diagnosis would not capture patients who developed Crohn's disease from Accutane exposure 366 days to four years after taking the drug. He noted, for example, that in the Crockett study the odds ratio increased from 0.68 (one-year analysis) to 0.89 (two-year analysis), which he said was likely a result of capturing more patients who had developed the disease. Similarly, the Bernstein study, which looked back approximately 2.6 years, found a positive association between Accutane and Crohn's disease, which Kornbluth opined may also have resulted from capturing more

Unrecognized Crohn's Disease And Celiac Disease, But Not Ulcerative Colitis, 56 Dig. Dis. Sci. 3270 (Nov. 2011).

²² V. Chouraki et al., The changing pattern of Crohn's disease incidence in northern France: a continuing increase in the 10- to 19-year-old age bracket (1988-2007), 33 Aliment. Pharmacol. Ther. 1133 (2011). Chouraki was not a study of the prodrome for Crohn's disease. It was a study of the increased incidence of the disease in northern France.

Crohn's disease patients than the other shorter studies. He found that these four studies were not designed to accurately account for all of the Accutane patients who had developed Crohn's disease, thereby distorting the results. He opined that if the studies had been designed to account for the long prodrome, the results would have shown a greater increased risk of developing Crohn's disease.²³

2. Power

Next, Kornbluth opined that three of the studies (Bernstein, Crockett, and Etminan), were insufficiently powered to detect a statistically significant association; in other words, that the sample size was not large enough to make a definitive conclusion as to whether there was a statistically significant risk. For example, in the Bernstein study, a large case-control study using a Canadian database, the study population only comprised 1118 Crohn's disease cases out of a control population of 19,419. He calculated that the "power" of the Bernstein study was low (about 25% to 30%), that is, there was only a 25% to 30% chance of

²³ Kornbluth's view about the length of the prodrome was hardly unique to him or scientifically unorthodox. It was supported by the Pimentel and Barratt studies, references in Passier, supra note 17, at 52, and admissions made by defendant's expert, Dr. Oliva-Hemker.

detecting a statistically significant association (greater than 2%) between Accutane and Crohn's disease. He explained that 80% power was appropriate for a study. He opined that if these studies had not been underpowered, the results would have been statistically significant for an increased risk of developing the disease.

However, Kornbluth admitted that where individual studies are underpowered to detect outcomes, studies can be pooled using a meta-analysis, to increase the power to detect a risk. One such study was done by Etminan, which combined the Bernstein, Crockett, Etminan (case-control study) and Racine studies, and found a pooled RR of .75 with a CI of 0.46 to 1.24, indicating no increased risk of developing Crohn's disease. Another study was done by Goodman, defendants' expert, as discussed more fully infra, who found a pooled RR of 0.87 with a CI of 0.59 to 1.28, or, again, no statistically significant increased risk of developing Crohn's disease. However, Kornbluth, like Madigan, rejected the results of these meta-analyses, explaining that a meta-analysis using underpowered and flawed studies (as he said existed in this case), which did not account for the prodrome, are not informative and should not be relied upon by scientists in determining causation.

3. Design flaws

Kornbluth also found the results of the five studies were inconclusive because they had design flaws, including differences in the populations and the failure to account for confounding variables. For example, he opined that the Bernstein study, a Canadian study, was flawed because of the differences in recommended doses between the United States (higher dose) and Canada (lower dose). He expected that given that difference there would be far fewer cases of Crohn's disease in Canada thereby decreasing the relative risk ratio. Similarly, there were differences in recommended doses between the United States (higher) and France (lower), which Kornbluth testified could account for the Racine study's finding of a protective effect.

Next, Kornbluth found that most of the studies had failed to account for smoking and family history confounders, that is, alternative causes of Crohn's disease unrelated to Accutane exposure that can bias the study by making the association appear higher or lower than it actually is.²⁴ For example, the Alhusayen

²⁴ Many of the studies had accounted for other confounders, including gender, oral contraceptive use, use of NSAIDs, and antibiotics. And the Etminan study adjusted, in part, for patients who had made a claim for tobacco cessation counseling.

study, which comprised 46,922 patients treated with Accutane, reported an unadjusted RR for Crohn's disease of 1.40, and an adjusted (for antibiotic use) RR of 1.17. He testified that the study should have adjusted for more relevant confounders, including family history and smoking, which presumably would have yielded a higher risk ratio of developing Crohn's disease from Accutane. In other words, removing all of the individuals who had a family history of Crohn's disease from the sample would yield a better measure of whether Accutane use is associated with Crohn's disease.

Lastly, Kornbluth did not consider the results of two additional studies (Rashtak and Fenerty), because those studies were designed to only report general IBD results, and did not calculate the relative risk of developing Crohn's disease, which has different triggers than ulcerative colitis and therefore did not provide reliable information on the risk of developing Crohn's disease.

d. Causality assessments

Next, Kornbluth considered defendants' internal causality assessments of adverse drug experience (ADE) reports of Crohn's disease in patients taking Accutane, in which defendants had

concluded that there was an association between Accutane and the disease. For example, in an internal causality assessment dated December 17, 2002, defendants reported that there were 159 reports of adverse events from exposure to Accutane received from worldwide sources; of those patients, sixty-four had Crohn's disease, of which Roche assessed causality as "related" in twenty-seven cases, with the remainder designated either as unrelated or unknown. Additionally, defendants concluded in an internal report dated November 16, 2000, that "[i]sotretinoin has been found to be causally associated with inflammatory bowel disease, including colitis."

Similarly, in its "general data memo," a document that reflected the company's scientific and medical opinion about Accutane, defendants provided that IBD "is a possible side effect of ROACCUTANE in very rare cases, possibly in patients predisposed to inflammatory gastro-intestinal diseases," and that although the side-effect "does not seem to represent a serious problem in practice," it is "reasonable to conclude" that the drug is "basically contraindicated" for patients "in the active phase" of IBD. Kornbluth testified that those "strong statement[s]" by defendants, who had a great deal of pharmacovigilance experience

in assessing ADE reports, were significant in assessing whether Accutane use was associated with Crohn's disease. In its core data sheet, which is a compilation of information set forth in labels or in reports sent to regulatory boards, defendants also stated that IBD had been reported in Accutane users and that adverse events are dose-related.

Existence of a Causal Relationship

After determining that there was an association between Accutane and Crohn's disease based on the literature, MedWatch reports, epidemiological studies, and causality assessments, Kornbluth then considered whether that association reflected a causal relationship.

a. Bradford Hill

As set forth above, epidemiology cannot prove causation. Reference Manual, supra, at 598. In making the causation determination, Kornbluth considered, among other factors, the widely recognized criteria identified by Sir Austin Bradford Hill (Bradford Hill criteria): (1) strength of the association; (2) temporal relationship; (3) consistency of relationship; (4) biological plausibility; (5) consideration of alternative explanations; (6) specificity; (7) dose-response relationship; (8)

replication; and (9) cessation of exposure. Sir Arthur Bradford Hill, The Environment and Disease: Association or Causation, 58 Proc. Royal Soc'y of Med. 295, 299 (1965). In assessing these criteria

[t]here is no formula or algorithm that can be used to assess whether a causal inference is appropriate based on these guidelines. One or more factors may be absent even when a true causal relationship exists. Similarly, the existence of some factors does not ensure that a causal relationship exists. Drawing causal inferences after finding an association and considering these factors requires judgment and searching analysis, based on biology, of why a factor or factors may be absent despite a causal relationship, and vice versa. Although the drawing of causal inferences is informed by scientific expertise, it is not a determination that is made by using an objective or algorithmic methodology.

[Reference Manual, supra, at 600.]

1. Strength of association

Kornbluth opined that the association evidence (scientific literature, MedWatch reports, epidemiological studies and causality assessments) reflected a strong association between Crohn's disease and Accutane, even though no medical organization or epidemiological study had concluded that Accutane causes Crohn's disease. He found that although the epidemiological studies had "some major shortcomings," most of the studies

nonetheless showed a "fairly substantial increased risk of Accutane causing Crohn's disease."

2. Temporal relationship

Kornbluth found that the medical literature, MedWatch reports, and animal studies supported a finding that there was a temporal relationship between Accutane exposure and Crohn's disease, that is, the timing of the exposure to the drug and the onset of the disease was consistent with the lengthy latency and prodromal period for the disease.

3. Biological plausibility

Although the precise mechanism by which Accutane could cause Crohn's disease is unknown, Kornbluth opined that there was a biologically plausible mechanism by which Accutane could cause Crohn's disease, namely, that Accutane may cause the migration of inflammatory T cells to the intestinal tract. Kornbluth explained that retinoic acid, which is a metabolite of Accutane, causes and perpetuates Crohn's disease by directing inflammatory T cells, using "antenna" known as "alpha 4 beta 7," to the intestines and allowing the T cells to bind to "receptors" (or "mucosal addressing cell adhesion molecules" ("MAdCAMs")), which then spurs invasion of inflammatory cells into the lining of the intestines. Without

retinoic acid, the antenna (alpha 4 beta 7), does not imprint on the T cells and are not guided back to the intestines. He said that studies have shown that blocking retinoic acid prevents intestinal inflammation, which is characteristic of Crohn's disease. In other words, without alpha 4 beta 7, one cannot "get Crohn's disease, because the T cells that are the driver of the inflammation have no way of getting into the small intestine."²⁵

In fact, he noted that two new drugs (Vedolizumab and Natalizumab), which Kornbluth said had been very effective in treating Crohn's disease, operated to block the retinoic acid antenna (alpha 4 beta 7), thereby preventing the T cells from binding to the MAdCAMs and entering the intestine where they cause damage. The success of these drugs indicated to Kornbluth "that retinoic acid is a damaging toxic pathway for patients with Crohn's disease," because inhibition of the harmful molecule caused the patient to get better.

Moreover, he noted that a Canadian case-control epidemiological study on children supported his opinion on biological plausibility because it reported that children

²⁵ Kornbluth's explanation as to biological plausibility had some support in scientific literature. See, e.g., Spada, supra note 17, at 24; Shale, supra note 17, at 737-39.

ingesting dietary supplements of retinol (vitamin A), a compound related to Accutane, which also breaks down into retinoic acid, had a statistically significant (two-fold) increased risk of developing Crohn's disease.²⁶ There was a dose effect, in that only the children taking higher than normal doses of retinol showed an increased risk of developing Crohn's disease.

4. Dose relationship

Next, Kornbluth testified that there was a dose-related relationship between Accutane and gastrointestinal injury—higher doses cause greater injury—as set forth in the Core Data sheet, dog studies, MedWatch reports, and epidemiological studies. He explained that "a dose-response curve" is "scientific evidence" of causation.

5. Consistency, coherence, and specificity

Lastly, Kornbluth testified that he had observed consistency across different lines of evidence supporting a causal relationship, including the MedWatch reports, dog studies, and the epidemiological studies, which except for the Racine study,

²⁶ Devendra K. Amre et al., Imbalances in Dietary Consumption of Fatty Acids, Vegetables, and Fruits Are Associated with Risk for Crohn's Disease in Children, 102 Am. J. Gastroenterol. 2016 (Sept. 2007).

reported an increased risk of developing the disease.

He also found that the evidence was coherent with the scientific understanding of the cause and presentation of the disease. For example, evidence from the dog studies was consistent with the knowledge of the pathogenesis of the disease in that a breakdown of the epithelium (as observed in some of the dogs) can serve as a trigger for Crohn's disease. The Vedolizumab and Natalizumab studies demonstrated that blocking the effect of retinoic acid (an Accutane metabolite) vastly improved a patient's Crohn's disease.

Kornbluth did not, however, find any specificity for Crohn's disease because the disease is not caused solely by Accutane use, and it is not the only side-effect of taking the drug.

b. Other evidence of a causal relationship

1. Animal studies

In reaching his conclusion on causation, Kornbluth relied on studies in which dogs were given high doses of Accutane (the dogs achieved similar levels of the active metabolite as humans because a dogs' metabolism is different). The studies reported gastrointestinal upset, diarrhea, bloody mucoid stools, intestinal adhesions, thickening of the mucosa, and epithelial damage, with

crypt abscess formation as seen in Crohn's patients, in the treated dogs. He explained that even though dogs cannot develop IBD, the studies showed that Accutane can cause "significant obvious symptomatic damage to the gastrointestinal tract" because a dog's intestine "is quite analogous to the human intestinal tract."

2. Challenge/dechallenge/rechallenge reports

Kornbluth opined that the challenge/dechallenge/rechallenge reports contained in the medical literature (Martin, Deplaix, and Melki) and in the MedWatch reports were "very compelling" evidence of causation. He explained that the reports of positive rechallenges were significant because they were essentially a non-deliberate human experiment, in that a potentially toxic substance was reintroduced to a patient resulting in further injury.

3. Class effects

Kornbluth also reviewed side effects reported from use of Vesanoid, a chemically similar retinoid manufactured by Roche used to treat acute promyelocytic leukemia (APL). Chemically, Vesanoid is tretinoin, an all-trans retinoic acid. Accutane, or isotretinoin, another retinoid, metabolizes into tretinoin and 4-oxo-isotretinoin. The Vesanoid package insert indicates that gastrointestinal disorders, including gastrointestinal

hemorrhage, were reported in thirty-four percent of clinical trial patients. These results were significant because such a high percentage of gastrointestinal disorders would not be expected, even in APL patients, who have a "tremendous tendency to bleed." Thus, he testified that this evidence supported his opinion that Accutane can cause gastrointestinal injuries.

2. Dr. Maria Oliva-Hemker

Like Dr. Kornbluth, defendant's gastroenterology expert, Dr. Oliva-Hemker, was highly qualified. She was board certified in gastroenterology and was a professor of medicine at Johns Hopkins University. She had treated hundreds of children who suffered from IBD, published more than seventy peer-reviewed scientific articles on IBD, and chaired various gastroenterology committees.

She opined that the available scientific evidence did not support a finding that Accutane can cause Crohn's disease. She testified that the scientific evidence supported a finding that retinoic acid had an anti-inflammatory or protective effect on the gastrointestinal tract. Moreover, she testified that all of the epidemiological data--the best available evidence--reported no increased risk of Crohn's disease associated with Accutane, which was consistent with the biological evidence of a protective effect.

She testified that Kornbluth had failed to follow well-recognized principles of medical evidence hierarchy by relying on lower-level data (such as case reports and animal studies), instead of higher-level epidemiological evidence.

However, on cross-examination, Oliva-Hemker admitted that Crohn's disease often has a lengthy prodrome. In fact, she admitted that information was reflected in her own professional writings.

B. Biostatistical and epidemiological experts

1. Dr. David Madigan

Plaintiff's expert, Dr. Madigan, had a Ph.D. in statistics, taught statistics at Columbia University, had published more than 150 papers on biostatistics and pharmacovigilance, and served as an investigator on an FDA pilot program to monitor the safety of FDA-regulated medical products. He did not testify as to causation, but rather explained the process of conducting epidemiological studies, and examined the six epidemiological studies on Accutane and Crohn's disease (Crockett, Bernstein, Alhusayen, Etminan, Racine, and Sivaraman).

Madigan was critical of the design of the epidemiological studies. He found that the studies were biased toward a finding

of decreased risk as a result of: 1) power (Crockett, Alhusayen, and Etminan); 2) prodrome (Crockett, Etminan, Racine, and Alhusayen); and 3) unmeasured confounders (Bernstein (dose and duration), Crockett (exposure and outcome), Alhusayen (allowing reentry of patients after 12-month period), Etminan (confined to contraceptive users), Racine (dose), and Sivaraman (dose and duration)).

For example, he found that the power of the studies, or "the power to detect a true effect of a particular size," was low. In reaching that determination, he employed standard statistical techniques and determined that the Accutane studies were not sufficiently powered to detect even a 50% increased risk -- a "meaningful" measure of risk. He calculated the nominal power of the four studies (that did not find a statistically significant risk to detect a 50% increased risk of Crohn's disease) as follows: Bernstein (37.8%); Crockett (18.2%); Alhusayen (89.4%); and Etminan (22.6%). In other words, the Bernstein study had only a 37.8% chance of finding a statistically significant increase when there is a 50% increased risk, or a 62.2% chance of finding a false negative.

He also opined that because Crohn's disease has a long

variable prodrome (ranging from a few months to several years), these studies, which focused on a short observation window to measure exposure, failed to account for all of the patients who developed Crohn's disease as a result of ingesting Accutane. Madigan explained that failing to account for just a few patients will "introduce bias into the study" toward a showing of no or decreased risk and will decrease the power of the study. Accounting for a 6.9-year prodrome (which he derived from the Pimentel study), Madigan calculated that the power of some of the studies further decreased to: Crockett (5.12%); Alhusayen (36.2%); and Etminan (4.5%). He found that only Bernstein (with its 2.6-year study) and Sivaraman (with its questionnaire format) had addressed prodrome through their study designs.

Moreover, Madigan testified, that it was not scientifically "appropriate to conduct a meta-analysis in this context, because of concerns with the individual studies." He explained that the purpose of a meta-analysis is to combine the results of epidemiological studies "to make a combined estimate," however, he noted that you cannot "make the bias" in a study "disappear" by combining several biased studies. Madigan, like Kornbluth, found that all of the studies in this case were "biased towards

the null," that is, "systematically biased to lower the estimated effect" and thus combining the studies would not yield an accurate result.

Instead, Madigan conducted a statistical disproportionality analysis of the spontaneous ADE reports for Accutane and Crohn's disease contained in the FDA's ADE reporting system database. He explained that spontaneous ADE reports "serve as a primary data source with which we, as a society, study drug safety concerns," and is an important component of drug safety investigation even though the data has limitations due to its reliance on voluntary reporting. He said that a disproportionality analysis is "standard" in analyzing ADE reports and is routinely used by the FDA and the pharmaceutical industry in assessing emerging safety concerns.

In his disproportionality analysis, which he conducted using the same methods as he used in conducting an analysis for the pharmaceutical industry, Madigan compared the observed rate of reporting for Accutane and Crohn's disease with the rate at which Crohn's disease was reported for other drugs in that database. He found that from 1997 to the present, there was a "striking signal of disproportionality" or a "strong association" between Accutane

use and Crohn's disease. Further, when Madigan removed the ADE reports generated through litigation by lawyer reporting, or approximately 88% of the reports, the results still showed a moderate increased risk of developing the disease.

He testified that a similar disproportionality analysis had been conducted, by researchers affiliated with the World Health Organization (WHO), of the WHO's drug safety database (Uppsala Monitoring Centre) in which the researchers compared the observed rate of ADE reports of Crohn's disease for Accutane to the observed rate of reports of Crohn's disease for other drugs in the WHO database.²⁷ They found a statistically significant increased risk (nineteen times greater) for developing Crohn's disease from Accutane use.

2. Dr. Steven N. Goodman

The defense expert, Dr. Goodman, had an M.D. degree, as well as a Ph.D. in epidemiology, was a professor and associate dean for clinical research at Stanford University, and had published numerous peer-reviewed scientific articles. He opined that the epidemiological evidence supported a finding that there was a

²⁷ The findings are cited in an article on case reports, Passier, supra note 17, at 52.

"strongly negative" association between Accutane and Crohn's disease, and that there was no biologic evidence or scientifically accepted causal mechanism that outweighed the results of the epidemiological studies. He opined that Kornbluth's and Madigan's methodology was not scientifically valid because they placed "very little weight" on the epidemiological studies, which were the highest tier of evidence in this case, and placed much more weight on "lesser forms of evidence," including case reports, animal studies, and causality assessments.

In forming his opinion, Goodman reviewed nine epidemiological studies (Bernstein, Crockett, Etminan (case-control and meta-analysis), Alhusayen, Racine, Rashtak, Sivaraman, and Fenerty). He considered the overall results of these studies for IBD as well as the results for Crohn's disease, because ulcerative colitis and Crohn's disease share a variety of risk factors and because it is difficult to distinguish between the diseases in the early stages. He opined that although the Sivaraman study showed an unadjusted statistically significant association between the drug and the disease, the study was too small to have any significance. He also opined that, viewed collectively, the epidemiological studies were consistent with Accutane having "either [a] potentially

protective effect" or "no effect."

To increase the power of the epidemiological studies, Goodman conducted a meta-analysis of Accutane and IBD (both ulcerative colitis and Crohn's disease), in which the larger more precise studies were given more weight. He found an RR of developing IBD of 0.87 (0.65-1.17), a negative association. He also conducted a meta-analysis of Accutane and Crohn's disease, and found a similar RR of 0.87 (0.59-1.28), another negative association.

Further, Goodman found that the epidemiological studies properly accounted for the prodrome of Crohn's disease. In determining the prodrome, Goodman reviewed several epidemiological studies, including the Chouraki study, and concluded that the average prodrome for Crohn's disease was nine months or less. He criticized Madigan's reliance on what Goodman characterized as the outlier non-population-based Pimentel study (6.9-years prodrome), which Goodman said was too small to provide any valid information. He concluded that the nine epidemiological studies, which applied a prodrome from one to two years, were properly designed and powered, and strongly supported a finding of no association between Accutane and Crohn's disease.

IV

Before we address plaintiffs' appellate arguments, it is helpful to review the legal principles applicable to the admissibility of expert testimony in toxic tort and similar cases.

To establish liability, plaintiffs must prove through expert testimony that ingestion of Accutane can cause Crohn's disease in humans (general causation). In addition, each individual plaintiff must prove specific causation, i.e., that Accutane was the cause of his or her disease. See DeLuca v. Merrell Dow Pharm., Inc., 911 F.2d 941, 958 (3d Cir. 1990). See also Perry v. Novartis Pharm. Corp., 564 F. Supp. 2d 452, 463 (E.D. Pa. 2008) ("Courts in toxic tort cases often separate the causation inquiry into general causation -- whether the substance is capable of causing the observed harm in general -- and specific causation -- whether the substance actually caused the harm a particular individual suffered."). The Kemp hearing at issue here concerned general, not specific, causation.

The admissibility of scientific evidence is governed by N.J.R.E. 702, which provides that "[i]f scientific, technical, or other specialized knowledge will assist the trier of fact to understand the evidence or to determine a fact in issue, a witness

qualified as an expert by knowledge, skill, experience, training, or education may testify thereto in the form of an opinion or otherwise." The Rule imposes three requirements:

(1) the intended testimony must concern a subject matter that is beyond the ken of the average juror; (2) the field testified to must be at a state of the art such that an expert's testimony could be sufficiently reliable; and (3) the witness must have sufficient expertise to offer the intended testimony.

[Hisenaj, supra, 194 N.J. at 15.]

The second requirement is at issue here, that is, whether Kornbluth's causation testimony and Madigan's statistical analysis testimony was sufficiently reliable in the field of scientific research to be admitted. Ibid.

In most cases, the proponent of scientific evidence must demonstrate that the opinions are "generally accepted, within the relevant scientific community" (the Frye standard). State v. Chun, 194 N.J. 54, 91, cert. denied, 555 U.S. 825, 129 S. Ct. 158, 172 L. Ed. 2d 41 (2008); State v. Harvey, 151 N.J. 117, 169-70 (1997) (citing Frye v. United States, 293 F. 1013, 1014 (D.C. Cir. 1923)), cert. denied, 528 U.S. 1085, 120 S. Ct. 811, 145 L. Ed.

2d 683 (2000).²⁸ See also Hisenaj, supra, 194 N.J. at 17. "That acceptance entails the strict application of the scientific method, which requires an extraordinarily high level of proof based on prolonged, controlled, consistent, and validated experience." Rubanick, supra, 125 N.J. at 436.

However, our Supreme Court has relaxed the "general acceptance" standard in tort cases involving injuries caused by toxic substances or medications, involving new or developing theories of causation. Kemp, supra, 174 N.J. at 430-31; Landrigan, supra, 127 N.J. at 414; Rubanick, supra, 125 N.J. at 449. Under the relaxed standard, "a scientific theory of causation that has not yet reached general acceptance may be found to be sufficiently reliable if it is based on a sound, adequately-founded scientific

²⁸ In 1993, the United States Supreme Court, construing the Federal Rules of Evidence, held that Federal Rule of Evidence 702 superseded Frye and mandated that the federal courts apply a more relaxed scientific reliability standard. Daubert v. Merrell Dow Pharm., Inc., 509 U.S. 579, 113 S. Ct. 2786, 125 L. Ed. 2d 469 (1993). Daubert was a pharmacological tort case involving the drug Benedictin. Id. at 582, 113 S. Ct. at 2791, 125 L. Ed. 2d at 476. In criminal cases, New Jersey courts have not followed Daubert, but continue to strictly apply the Frye test, i.e., whether the scientific community generally accepts the reliability of the proffered evidence. See Harvey, supra, 151 N.J. at 168. As noted, in civil cases involving toxic torts, our courts use the relaxed standard set forth in Rubanick. See Kemp, supra, 174 N.J. at 430-31.

methodology involving data and information of the type reasonably relied on by experts in the scientific field." Rubanick, supra, 125 N.J. at 449. Thus, the Court "changed the focus of the inquiry from the scientific community's acceptance of the substance of the opinion to its acceptance of the methodology and reasoning underlying it." Clark v. Safety-Kleen Corp., 179 N.J. 318, 337 (2004). The Rubanick standard does not require the "extraordinarily high level of proof[,]" ordinarily required before a scientific theory will attain general acceptance in the scientific community. Rubanick, supra, 125 N.J. at 436.

The rationale behind relaxation of the standard was "the extraordinary and unique burdens facing plaintiffs who seek to prove causation" in such cases. Id. at 433.²⁹ The task of proving causation in toxic tort cases "is invariably made more complex because of the long latency period of illnesses caused by carcinogens or other toxic chemicals." Ayers v. Jackson, 106 N.J. 557, 585 (1987). And, in some drug cases, causation may not have

²⁹ Rubanick relied heavily on persuasive federal court opinions, in rejecting the general acceptance test. Rubanick, supra, 125 N.J. at 445 (citing United States v. Downing, 753 F.2d 1224, 1237 (3d Cir. 1985), and DeLuca, supra, 911 F.2d at 941). See also Ferebee v. Chevron Chem. Co., 736 F.2d 1529 (D.C. Cir.), cert. denied, 469 U.S. 1062, 105 S. Ct. 545, 83 L. Ed. 2d 432 (1984) (cited in Rubanick, supra, 125 N.J. at 440).

been "confirmed by the scientific community but compelling evidence nevertheless suggests that such a relationship exists." Kemp, supra, 174 N.J. at 430.

V

Against that legal backdrop, we consider plaintiffs' contention that the trial court erred in barring their experts' testimony. Plaintiffs assert that: 1) their experts relied on methodologies and data of the type relied on by comparable experts; 2) the judge substituted his judgment on the epidemiological studies for that of the expert scientists; 3) the experts' reliance on the studies involved a methodology generally followed by comparable experts; 4) Kornbluth appropriately considered the epidemiological studies in assessing the relationship between Accutane and Crohn's disease; 5) the judge mischaracterized Madigan's testimony; 6) the judge impermissibly weighed evidence of ADE reports and animal studies; and 7) the judge abused his discretion in assessing the credibility of plaintiffs' experts. We agree that the judge erred in excluding the experts' testimony.

Under the relaxed standard, as applicable here, the trial court assesses "the soundness of the proffered methodology and the qualifications of the expert." Kemp, supra, 174 N.J. at 426

(quoting Rubanick, supra, 125 N.J. at 454). The focus of the trial court's inquiry must be "solely on principles and methodology, not on the conclusions that they generate." Ibid. (quoting Daubert, supra, 509 U.S. at 594-95, 113 S. Ct. at 2797, 125 L. Ed. 2d at 484). In evaluating the methodology, "courts should consider whether others in the field use similar methodologies." Rubanick, supra, 125 N.J. at 449.

In making that determination the court should not "directly and independently determine as a matter of law that a controversial and complex scientific methodology is sound." Id. at 451. "The critical determination is whether comparable experts accept the soundness of the methodology, including the reasonableness of relying on this type of underlying data and information. Great difficulties can arise when judges, assuming the role of scientist, attempt to assess the validity of a complex scientific methodology." Ibid. Nor is it appropriate for the trial judge to second-guess an expert's interpretation of the underlying data. Ibid. For example, in Rubanick the Court found that:

the trial court . . . "independently reviewed" each of the thirteen studies on which Dr. Balis relied, and decided that they "do not say what plaintiff's expert concludes." In engaging in such an analysis, the court substituted its own assessment of the studies

for that of an acknowledged expert. . . .
"[t]he interpretation of the data . . . is the
function of the qualified expert
[C]ourts should be loath to determine whether
the particular expert has properly relied upon
data which experts in the field generally rely
on." Thus, the inquiry is not the reliability
of the expert's ultimate opinion nor is it
whether the expert thought his or her own
reliance on the underlying data was
reasonable, nor whether the court thinks that
the expert's reliance was reasonable[.] The
proper inquiry is whether comparable "experts
in the field [would] actually rely" on that
information.

[Id. at 451-52 (quoting Ryan v. KDI Sylvan Pools, Inc., 121 N.J. 276, 289 (1990) (additional citations omitted).]

The qualifications of the expert must also "be factored into the determination of the soundness of the methodology used." Id. at 452.

"If epidemiological studies are to provide the basis for an expert's opinion, they must have been 'soundly and reliably generated' and be 'of a type reasonably relied on by comparable experts in the particular field.'" Landrigan, supra, 127 N.J. at 419-20 (quoting Rubanick, supra, 125 N.J. at 447). When an expert relies on epidemiological studies, the "court should review the studies, as well as other information proffered by the parties, to determine if they are of a kind on which such experts ordinarily

rely." Id. at 417. Significantly, the court must "examine the manner in which experts reason from the studies and other information to a conclusion[,]" which "must derive from a sound methodology that is supported by some consensus of experts in the field." Id. at 420. See In re Zolof Prods. Liab. Litig., 26 F. Supp. 3d 449, 460 (E.D. Pa. 2014) (excluded expert whose "opinion regarding class effects is not evidence based, and is directly contrary to the findings of her own peer-reviewed, published research"); In re Rezulin Prods. Liab. Litig., 369 F. Supp. 2d 398, 425 (S.D.N.Y. 2005) (court excluded expert testimony where expert selectively chose his support from scientific literature and failed to "acknowledge or account for" evidence that tended to refute his theory).

A court's assessment of scientific expert evidence should include an evaluation of the studies upon which the experts rely, but the court must not substitute "its own assessment of the studies for that of an acknowledged expert." Rubanick, supra, 125 N.J. at 451. "Although trial courts are expected to act as gatekeepers to the proper admission of expert testimony," courts are not expected "to investigate sua sponte the extent to which the scientific community holds in esteem the particular analytical

writings or research that a proponent of testimony advances as foundational to an expert opinion." Hisenaj, supra, 194 N.J. at 16. "The court's function is to distinguish scientifically sound reasoning from that of the self-validating expert, who uses scientific terminology to present unsubstantiated personal beliefs." Landrigan, supra, 127 N.J. at 414. The plaintiff bears the burden of proof in establishing admissibility. Kemp, supra, 174 N.J. at 429.

Ordinarily, the admission or exclusion of expert testimony is "committed to the sound discretion of the trial court[,]" Townsend v. Pierre, 221 N.J. 36, 52 (2015), and we review the decision for abuse of discretion. Hisenaj, supra, 194 N.J. at 12. However, we owe somewhat less deference to a trial court's determination in a case of this type. See State v. Torres, 183 N.J. 554, 567 (2005). "Although 'the trial court is in a better position to shape the record and make credibility determinations,' an 'appellate court need not be as deferential to the trial court's ruling on the admissibility of expert scientific evidence as it should be with the admissibility of other forms of evidence.'" State v. J.R., 227 N.J. 393, 410 (2017) (quoting Torres, supra, 183 N.J. at 567). In determining whether the trial court

misapplied discretion, we consider whether the court's analysis of the evidence was faithful to the principles set forth in Rubanick, or whether the court misapplied the standards. See State v. Darby, 174 N.J. 509, 518 (2002) (no deference is to be accorded to the trial court's decision to admit other-crime evidence, nor is that decision entitled to be reviewed under an abuse of discretion standard, where the trial judge failed to apply applicable law); Konop v. Rosen, 425 N.J. Super. 391, 401 (App. Div. 2012) (appellate review is de novo when the trial court fails to apply the proper test in analyzing the admissibility of proffered evidence). See also Pressler & Verniero, Current N.J. Court Rules, comment 4.7 on R. 2:10-2 (2017) ("When the trial court fails to apply the proper test in analyzing the admissibility of proffered evidence, the de novo standard of review . . . applies").

Here, the court found that although both of plaintiffs' experts were "eminently qualified, their reasoning and methodology is slanted away from objective science and in the direction of advocacy." The court found that Kornbluth's methodology was not supported by the scientific community because he interpreted the Sivaraman study differently than its authors did. Similarly, the

court found that Madigan placed undue weight on the Sivaraman study, and "ignored" the other studies. The judge also criticized Madigan for failing to perform a meta-analysis of all of the studies.

The court further reasoned that plaintiffs' experts had failed to follow valid scientific methodology in relying on the Sivaraman and Pimentel studies to the exclusion of the larger population-based epidemiological studies, concluding that the "scientific literature does not support reliance upon such insignificant studies to arrive at conclusions." The court concluded that Kornbluth's "contrived reasoning is not supported by the scientific community as a reliable basis for making causal determinations," and Madigan's opinions were not methodology based, but rather conclusion-driven.

Additionally, the court found that Kornbluth's testimony was "replete with what can be described as convenient assumptions. When he needs to bridge an analytical gap in his methodology he assumes facts, events and conclusions as he wants them to be in support of his hypothesis." For example,

in response to counsel's questioning regarding the results of various studies, Dr. Kornbluth assumed: (a) that all the patients in the two studies upon which he relied filled out their

questionnaires correctly; (b) despite the fact that the authors of the Sivaraman study got it wrong as to their adjustment for antibiotics, he assumed they got everything else correct; (c) he assumed that in the Rashtak Study, the patients with Accutane exposure were followed for less time than the control group; and (d) he assumed the size of the doses of Accutane given to the subjects in various studies.

With regard to biological plausibility, the court found

Kornbluth's discussion of his hypothesis for the biological mechanism of the development of CD [Crohn's disease] as caused by Isotretinoin falls far short of being "compelling." His basis for the discussion are the medications Natalizumab and Vedolizumab. He attempts to extrapolate causation of CD by Isotretinoin by discussing treatment of CD by these other medications. Dr. Oliva-Hemker [sic] explained the inherent weakness of trying to rely upon the data on Natalizumab and Vedolizumab as being probative of causation. In essence, treating a "pathway" that develops once a disease occurs, does not mean that . . . a particular treatment mechanism informs as to the original cause of the disease. She also pointed out that this hypothesis is contrary to a significant body of scientific literature showing that Retinoic acid is actually anti-inflammatory . . .

The court described Madigan as "an expert on a mission," and criticized Kornbluth's approach as being less convincing than Oliva-Hemker's analysis as to causation. He was also critical of plaintiffs' experts' reliance on lines of evidence other than

epidemiological studies:

[C]oursing through Plaintiffs' presentation is a refrain that is a ruse. Repeatedly, counsel for the Plaintiffs and their witnesses spoke of "lines of evidence," emphasizing that their experts examined the same "lines of evidence" as did the experts for the Defense. Counsels' sophistry is belied by the fact that the examination of the "lines of evidence" by Plaintiffs' experts was highly selective, looking no further than they wanted to -- cherry picking the evidence -- in order to find support for their conclusion-driven testimony in support of a hypothesis made of disparate pieces, all at the bottom of the medical evidence hierarchy. This crafty stratagem cannot bridge the analytical gaps inherent in Plaintiffs' hypothesis.

Plaintiffs contend that, whether or not the trial judge found their experts opinions persuasive in substance, the experts relied on methodologies and data of the type reasonably relied upon by comparable experts. We agree.

A. Data and Information

Whether or not it persuades a jury, it is clear to us that in forming their conclusions Kornbluth and Madigan relied on the types of data and information reasonably relied on by comparable experts in the scientific field, and by the experts in previous Accutane cases in this docket. That evidence includes epidemiological studies, scientific articles, case studies,

clinical studies, animal studies, and causality assessments.

It is well-established that epidemiological studies, published in peer-reviewed scientific journals, as were considered in this case, are the type of data reasonably relied on by the scientific community to determine whether exposure to a drug is associated with a disease. Landrigan, supra, 127 N.J. at 419-20. Properly conducted epidemiological studies are a significant factor in establishing causation in toxic tort cases. See Reference Manual, supra, at 551 n.2. "[E]pidemiology is a well-established branch of science and medicine, and epidemiological evidence has been accepted in numerous cases." DeLuca, supra, 911 F.2d at 954. See Magistrini v. One Hour Martinizing Dry Cleaning, 180 F. Supp. 2d 584, 591 (D.N.J. 2002) (epidemiological studies), aff'd o.b., 68 Fed. App'x 356 (3d Cir. 2003). Notably, although epidemiological evidence is not required to prove causation, if it exists, an expert cannot ignore it. Perry, supra, 564 F. Supp. 2d at 465. However, the existence of inconclusive epidemiological studies does not preclude an expert from relying on alternative data, such as animal studies. Id. at 466.

Moreover, although the Sivaraman epidemiological study was published as an abstract and not a full article, and was presented

at a conference, our courts recognize that "[s]upport for an expert's methodology may be found in professional journals, texts, conferences, symposia, or judicial opinions accepting the methodology." Kemp, supra, 174 N.J. at 427. In accord with accepted scientific methodology, Kornbluth also considered other forms of evidence in determining causation, including animal studies, case reports, challenge/dechallenge/rechallenge reports, causality assessments, class effects, and published scientific literature. Although case reports and causality assessments should be interpreted with caution, there was nothing so inherently unreliable about the materials Kornbluth cited as to preclude their consideration as part of a scientific expert's methodology under N.J.R.E. 702. Further, the experts did not elevate this evidence over the epidemiological studies, but rather considered this evidence in forming their opinions.

B. Methodology and Reasoning

We also conclude that the methodology used by Kornbluth and Madigan to reach their conclusions was consistent with sound scientific principles and methodologies accepted in the medical and scientific community. Rubanick, supra, 125 N.J. at 449; Kemp, supra, 174 N.J. at 431. In making that determination the court

must "examine the manner in which experts reason from the studies and other information to a conclusion." Landrigan, supra, 127 N.J. at 420. The experts' conclusions "must derive from a sound methodology that is supported by some consensus of experts in the field." Ibid. The experts must identify the factual bases for their conclusions, explain their methodology, and demonstrate that both the factual bases and the methodology are reliable. Kemp, supra, 174 N.J. at 427; Rubanick, supra, 125 N.J. at 449-50.

The primary focus in this appeal is upon Kornbluth's and Madigan's analysis of the epidemiological studies. Those studies indicated that the relationship between Accutane and Crohn's disease is more tenuous than the relationship between the drug and ulcerative colitis. For example, Crockett (a large study) found a statistically significant association between Accutane and ulcerative colitis, and four studies (Bernstein, Etminan, Racine and Alhusayen) found a positive association between the drug and the disease.

In contrast, only one small study (Sivaraman) found a statistically significant positive association (before adjustment for antibiotic use) between Accutane and Crohn's disease, two found a positive association (Bernstein and Alhusayen), three

found a negative association (Crockett, Etminan, and Racine), and one (Racine) found a statistically significant association for a protective effect. The degree to which this contrary opinion dominates the epidemiological studies is relevant to the reliability inquiry. DeLuca, supra, 911 F.2d at 955. However, demonstrable flaws in the studies may undercut their significance. Ibid.

In other words, does the relevant scientific community accept the process by which Kornbluth and Madigan reasoned to a conclusion that the epidemiological studies (despite the lack of a statistically significant association) and the other relevant evidence supported a finding of a causal relationship between Accutane and Crohn's disease? In some cases, a court may conclude that there is simply too great an analytical gap between the data and the expert's opinion. See Gen. Elec. Co. v. Joiner, 522 U.S. 136, 146, 118 S. Ct. 512, 519, 139 L. Ed. 2d 508, 519 (1997). However, in this case, we conclude that the data was sufficient to permit the experts to testify, and any weaknesses in their opinions can be explored through cross-examination.

It is well-established that "[t]he usefulness of an epidemiological study depends on the quality of the underlying

data, the reliability of the methodology, and the validity of the interpretations." Landrigan, supra, 127 N.J. at 420 (quoting Michael Dore, A Commentary on the Use of Epidemiological Evidence in Demonstrating Cause-in-Fact, 7 Harv. Envtl. L. Rev. 429, 432 (1983)). An expert should, under sound scientific methodology, evaluate the study in assessing its validity. See ibid. We infer from the Manual that an expert should consider a study's possible flaws and weaknesses before deciding whether to rely on it. Reference Manual, supra, at 553. Further, an expert can rely on the data generated from a study even if he or she disagrees with the author's conclusion, and need not subject his or her own analysis to peer review. DeLuca, supra, 911 F.2d at 954.

Here, in contrast to the court's finding, Kornbluth and Madigan, in accordance with established scientific methodology, evaluated all of the epidemiological and prodrome studies, not just Sivaraman and Pimentel. They testified at length as to the design flaws and limitations of the epidemiological studies, including the failure to account for the prodrome, insufficient power, and design flaws, which are all recognized in the scientific community as capable of producing an erroneous association in an epidemiological study. See Reference Manual, supra, at 583. For

example, a poorly conceived or conducted study that disproves the null hypothesis at a high level of significance may be far less reliable than a well-conceived and conducted study that is significant at a lower level. DeLuca, supra, 911 F.2d at 955.³⁰ Moreover, the fact that defendants' experts interpret the epidemiological studies differently does not, standing alone, indicate that Kornbluth and Madigan failed to rely upon a sound methodology. "Indeed, 'in most cases, objections to the inadequacies of a study are more appropriately considered an objection going to the weight of the evidence rather than its admissibility.'" Rosenfeld v. Oceania Cruises, Inc., 654 F.3d 1190, 1193 (11th Cir. 2011) (quoting Hemmings v. Tidyman's Inc., 285 F.3d 1174, 1188 (9th Cir. 2002), cert. denied, 537 U.S. 1110, 123 S. Ct. 854, 154 L. Ed. 2d 781 (2003)).

That said, certainly, as the trial judge correctly observed,

³⁰ DeLuca was "a diversity action brought under New Jersey law" against the manufacturer of Benedictin. Id. at 942-43. Anticipating Daubert, the Third Circuit Court of Appeals rejected the Frye standard, and reversed a trial court decision barring testimony from an expert whose approach was remarkably similar to that of plaintiffs' experts in this case. Id. at 955. Although DeLuca is not binding on us, we find it persuasive. As previously noted, DeLuca was also one of the seminal federal cases cited with approval in Rubanick as being "compatible with our own rules of evidence." Rubanick, supra, 125 N.J. at 445, 447.

larger studies enable researchers to form a more accurate conclusion and reduce the chance of random error in their results. Reference Manual, supra, at 576. However, Kornbluth could, in applying accepted scientific methodology, properly consider one small well-designed study over larger seriously flawed studies as a basis for drawing an inference about the studied subject.³¹ Further, Kornbluth did not ignore the results of the larger studies in favor of the smaller Sivaraman study, but considered the relative risk and the bounds of the 95% confidence interval in reviewing the conclusions. He also found that although the results of the studies were inconclusive, they were informative on his theory of causation. For example, he noted that four of the studies (Bernstein, Etminan (unadjusted), Alhusayen, and Sivaraman) found a positive association between Accutane use and Crohn's disease, and that one of the studies (Bernstein) showed an increased association when accounting for a two-year but not a one-year prodrome.

³¹ Scientific acceptance of small studies is not unknown. A treatise on pharmacoepidemiology, which is included in the parties' appendices, notes that "using case control studies, one can study rare diseases with markedly smaller sample sizes. . . . For example, the classic study of diethylstilbestrol and clear cell adenocarcinoma required only 8 cases and 40 controls." (Brian L. Strom, Pharmacoepidemiology 23 (4th ed. 2006)).

Similarly, Kornbluth testified at length as to the results of both the Pimentel (6.9-year prodrome) and Barratt (four-year prodrome) prodrome studies and opined that the results of those studies were in agreement with his decades of experience treating thousands of Crohn's disease patients. Kornbluth and Madigan rejected, but did not ignore, the findings of the Chouraki study based on the selection of patients and the use of patient charts. Further, Madigan testified that the Pimentel study, which utilized questionnaires, contained the most useful and relevant data, including raw data from which he could compute age-specific estimated prodromes.³²

Kornbluth's methodology in analyzing the epidemiological studies was also bolstered by Madigan, who presented detailed testimony as to the insufficient power of the epidemiological studies. Further, Madigan testified that the decision whether to conduct a meta-analysis is a scientific judgment, and explained

³² Our reading of the Pimentel article supports Kornbluth's description of the authors' very meticulous approach to gathering and verifying information about the subjects. We cannot agree with the trial judge's view that Kornbluth was "cherry picking" in relying on Pimentel. Kornbluth cogently explained why he believed that the authors' methodology was reliable, and consistent with his own medical practice in diagnosing Crohn's disease patients.

that such an analysis would not yield reliable results in this case. Instead, he conducted and reviewed a disproportionality analysis, which while not without limitations, is a validated method in drug safety research and surveillance.

Both Kornbluth and Madigan explained in considerable detail why most of the studies were biased toward "the null" or no effect, and were otherwise inadequate to reliably demonstrate whether or not there was a statistically significant connection between Accutane and Crohn's disease. They also explained why the statistically significant initial results of the Sivaraman study were more reliable than the adjusted results.

The reliability of Kornbluth's opinion on causation was also strengthened by his consideration of other evidence (including case reports, animal studies, and causality assessments), and most notably, because he presented a biologically plausible mechanism for how Accutane causes Crohn's disease. See Reference Manual, supra, at 604 ("When biological plausibility exists, it lends credence to an inference of causality"). Once Kornbluth found that there was an association between Accutane and Crohn's disease based on his reading of the epidemiological studies, in addition to the scientific articles, MedWatch reports, and causality

assessments, he then considered whether that association was causal, utilizing the Bradford Hill criteria. Under that analysis, he considered, among other criteria, whether there was a biologically plausible mechanism by which Accutane could cause Crohn's disease—an important factor for determining a causal relationship.

Kornbluth, like plaintiffs' causation expert in previous Accutane trials, presented a biologically plausible mechanism supported by scientifically authoritative sources. He opined, based on his experience as a board-certified gastroenterologist and in conducting clinical trials on several drugs intended for use in the management of IBD, that retinoic acid, a metabolite of Accutane, "is a damaging pathway for patients with Crohn's disease." He found support for that opinion in the fact that two new drugs (Vedolizumab and Natalizumab) were effective in treating Crohn's disease and in a Canadian case-control epidemiological study that reported an increased risk of Crohn's disease from retinoic acid. Olivia-Hemker disagreed with that opinion and cited to other studies that supported a finding that retinoic acid had an anti-inflammatory or protective effect on the intestines, but that dispute goes to the weight, not the admissibility of the

testimony. See Harvey, supra, 151 N.J. at 178.

In conclusion, Kornbluth and Madigan, who are indisputably extremely well-qualified experts, considered all of the relevant data and information, applied appropriate methodology in analyzing the epidemiological studies, and expressed valid reasons for rejecting the conclusions of some of the epidemiological studies and in accepting other studies as supportive of their opinion. Although the relationship between the epidemiological scientific evidence and the experts' opinions is more tenuous than the evidence as to ulcerative colitis, the studies do not render the experts' testimony inadmissible. The manner in which plaintiffs' experts reasoned from the results of the epidemiological studies and other data is sufficiently sound to be reliable. Landrigan, supra, 127 N.J. at 420. Further, the experts did not ignore the findings of the larger epidemiological studies but explained the scientific bases for their criticism of the studies. Defendants' criticisms of the experts' choices as to the evidence on which they relied, can be addressed during cross-examination at trial. Hisenaj, supra, 194 N.J. at 24.

We also cannot agree with the trial court's view that because Kornbluth had not submitted his "current hypothesis" to a peer-

reviewed publication, he must have generated his opinion solely as a result of litigation and was a mere "hired gun." An expert is not required to submit her own analysis to peer review in order for a court to consider it. See DeLuca, supra, 911 F.2d at 954. We also do not subscribe to the trial court's characterization of Madigan as a hired gun whose testimony "was needed to clear the way for Dr. Kornbluth's hypothesis and that was the role he played, without regard to whether or not his efforts led the discussion any closer to scientific truth." Madigan carefully explained his methodology and his testimony should not have been discounted because defendants heavily contested his conclusions.

Given that our evidence rules embody a strong preference for admissibility, we conclude that the court mistakenly applied its discretion in excluding the expert scientific testimony. See N.J.R.E. 702; N.J.R.E. 401; State v. Jenewicz, 193 N.J. 440, 454 (2008) (noting "Rule 702's tilt in favor of the admissibility of expert testimony"); State v. Granskie, 433 N.J. Super. 44, 47-48 (App. Div. 2013); Kuehn v. Pub Zone, 364 N.J. Super. 301, 320 (App. Div. 2003) (expert's testimony was relevant to issues under consideration), certif. denied, 178 N.J. 454 (2004).

VI

In concluding, we emphasize the following observations. The trial court's decision, and our decision of this appeal, must be viewed in the context of this particular MCL litigation. It presents a close question concerning the survival of plaintiffs' cause of action in the face of new scientific information about Accutane and IBD.

In deciding this appeal, we bear in mind that science is constantly evolving, and that under our State's legal precedents, legal decision making in toxic tort and similar cases may vary from scientific decision making. The opportunity of thousands of plaintiffs, claiming injury from Accutane, to have their day in court may rest on that difference and must be decided now.

In general . . . clinical, regulatory, commercial, and legal decisions need to be made based on the best evidence available at the time of the decision. To quote Sir Austin Bradford Hill:

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.

Who knows, asked Robert Browning, but the world may end tonight? True, but on available evidence, most of us make ready to commute on the 8:30 next day.

[Pharmacoepidemiology, supra, at 26-27 (quoting Hill, supra, at 295-300).]

The parties in this case differ sharply on the question of what constitutes "the best evidence available at the time of the decision." Id. at 26. In particular, the case presents the question whether, in the face of several epidemiological studies that do not demonstrate a statistically significant relationship between taking Accutane and developing Crohn's disease, plaintiffs can continue to rely on other types of evidence - which in this same MCL docket they were previously permitted to use - to prove general causation. We also consider whether they can rely in part on information from some of the epidemiological studies that show a positive correlation, albeit not reaching the level of statistical significance.

We conclude that in this case, the epidemiology studies are not a conclusive bar to plaintiffs' case, and that their experts should be allowed to testify. Although epidemiological studies are considered as high on the tier of evidence bearing on the

question of causation, like any other form of scientific evidence, any particular study is only valuable if it is conducted in a scientifically reliable manner. Any party - plaintiff or defendant - has the right to challenge the methodology and, hence the results, of an epidemiological study.

In fact, the Reference Manual on Scientific Evidence cautions that

[A]ll [epidemiological] studies have "flaws" in the sense of limitations that add uncertainty about the proper interpretation of the results. Some flaws are inevitable given the limits of technology, resources, the ability and willingness of persons to participate in a study, and ethical constraints. In evaluating epidemiologic evidence, the key questions, then, are the extent to which a study's limitations compromise its findings and permit inferences about causation.

[Reference Manual, supra, at 553.]

After explaining some of the most common biases that may affect observational epidemiological studies, the Manual states that "[t]here are dozens of other potential biases that can occur in observational studies, which is an important reason why clinical studies (when ethical) are often preferable." Id. at 590. Thus it can be expected that, as in this case, the methodology and limitations of epidemiological studies - particularly

observational studies - will be fertile ground for disagreement among experts.

Moreover, the Manual supports plaintiffs' continued reliance on other types of evidence to prove their case, particularly given their well-explained opinions that most of the epidemiological studies are fundamentally flawed. Contrary to the trial judge's view, the Manual does not discount the value of live animal studies. While noting some potential weaknesses of the studies, the Manual states that toxicological studies, of which animal studies are one type, are "often . . . the only or best available evidence of toxicity." Id. at 564. The Manual also cautions that "[w]here both animal toxicologic and epidemiologic studies are available, no universal rules exist for how to interpret or reconcile them. Careful assessment of the methodological validity and power of the epidemiologic evidence must be undertaken, and the quality of the toxicologic studies and the questions of interspecies extrapolation and dose-response relationship must be considered." Id. at 564-65.

The judge, and defendants, relied heavily on a section of the Manual captioned "Hierarchy of medical evidence" (the medical hierarchy section). Id. at 723. However, that section does not

appear in the Reference Guide on Epidemiology. Rather, the section appears in the Reference Guide on Medical Testimony, as part of a chapter on medical decision-making. That chapter describes how doctors make decisions about diagnosing and treating patients and discusses the difficulties they face in making those decisions. Id. at 704. We do not construe the medical hierarchy section of the Manual as prescribing a rigid hierarchy for the acceptance or rejection of evidence in a legal setting. See Matrixx Initiatives, Inc. v. Siracusano, 563 U.S. 27, 40-42, 131 S. Ct. 1309, 1318-20, 179 L. Ed. 2d 398, 410-12 (2011); DeLuca, supra, 911 F.2d at 957. In fact, the preface to the Manual cautions judges as to "the proper use of the reference guides. They are not intended to instruct judges concerning what evidence should be admissible or to establish minimum standards for acceptable scientific testimony." Reference Manual, supra, at xv. As significantly, nothing in the Manual suggests that once epidemiological studies have been done, they are beyond scientific criticism, and no countervailing evidence should be considered.

We cannot agree with the trial judge's observation that plaintiff's experts "ignored" the epidemiological studies in favor of less reliable evidence. The experts did not ignore the studies.

Rather, in extensive and detailed testimony, they opined that most of the studies were unreliable, and they explained in considerable detail the reasons for those opinions.

In their testimony, both of plaintiffs' experts raised fundamental objections to the way the studies were conducted - particularly the length of time for which the studies followed the subjects. Based both on a prodrome study he found reliable and on the decades he has spent treating thousands of Crohn's patients, Dr. Kornbluth testified that the prodrome for Crohn's disease is much longer than the one-year time frame covered by most of the studies. Defendant's biostatistical expert, Dr. Goodman, admitted that if Dr. Kornbluth was correct about the prodrome, then all of the epidemiological studies on which the defense relied would be flawed.

Kornbluth also explained in detail other weaknesses of several of the studies. For example, the Alhusayan study treated subjects exposed to Accutane as being non-exposed after a one-year period following treatment. Thus, if those subjects developed IBD after a year and a day, the study reported them as though they had

never taken Accutane.³³ Kornbluth also explained that studies from other countries would not necessarily reflect the experience of United States subjects, because the standard dose of Accutane given to patients in those other countries is half that given to patients in the United States.

Kornbluth's view on the prodrome issue was actually bolstered by some of the defense testimony. During her cross-examination, Dr. Oliva-Hemker was confronted with her own book, which answered the question "How long have I had my IBD" by advising that: "Some people have years of symptoms before the diagnosis [of IBD] is made, while in others, these symptoms appear suddenly. Both groups may have had intestinal inflammation for days, months, or years, even though they didn't experience any symptoms at all for most of that time." She then clarified that "we traditionally apply that more to Crohn's patients rather than ulcerative colitis patients in terms of [it taking] years" to diagnose the disease. She was also confronted with a book written by a recognized expert who referred to "the four-year average delay of diagnosis of

³³ Although the study authors downplayed the results, the Alhusayan study also discovered what the authors characterized as a "weak" but statistically significant connection between Accutane and the development of IBD in teenagers, ages twelve to nineteen.

Crohn's disease."

Oliva-Hemker also confirmed that IBD affects about one percent of the population in the United States; Crohn's disease is a small subset of IBD, so the proportion of persons with Crohn's is much smaller than one percent. Those admissions support the view of plaintiff's experts that a study's failure to detect even a small number of "exposed cases," i.e., persons with Crohn's disease who had taken Accutane, could produce skewed results.

Further, when it suits their litigation strategy, defendants do not treat epidemiological studies as the last word in scientific proof. During the cross-examination of Oliva-Hemker, she admitted that in earlier Accutane litigation, when four epidemiological studies - concerning the lack of connection between antibiotics and ulcerative colitis - did not support her opinion that the plaintiff's UC was caused by antibiotics rather than Accutane, she relied on evidence of biological plausibility instead. It took almost four pages of repetitive questioning before Oliva-Hemker finally admitted that the methodology she used was valid. Yet, in this litigation, she criticized Kornbluth for relying on evidence of biological plausibility and placing less weight on the epidemiological studies.

Additionally, during the cross-examination of defendant's epidemiology expert, Dr. Goodman, he admitted that some of the epidemiological studies in this case had biases and weaknesses. He admitted, for example, that none of the studies controlled for family history, even though that is recognized as a strong factor in a person's potentially developing Crohn's disease. He contended that scientific judgment was required to evaluate how important those biases or weaknesses were. Goodman was also confronted with some of his own writings, in which he stated that, "If bias is present in each or some of the individual studies, meta-analysis will simply compound the errors and produce a wrong result that may be interpreted as having more credibility." That same point was made by plaintiff's epidemiology expert, Dr. Madigan, and it finds support in the Manual.

We appreciate that the trial judge had the opportunity, which we did not, to see the witnesses testify firsthand. However, his extreme negative reaction to plaintiffs' witnesses is not supported by the trial record. See J.R., supra, 227 N.J. at 410; Torres, supra, 183 N.J. at 567. In reviewing Madigan's testimony, we cannot agree with the judge that Madigan was a biased expert "on a mission." His testimony was coherent and consistent, and the

attorney who cross-examined him made little headway in discrediting his direct testimony.

The judge's disapproval of plaintiffs' experts' reliance on "lines of evidence" seems misplaced, because the defense used the same terminology and considered the same evidence. Dr. Oliva-Hemker agreed that she and Dr. Kornbluth looked at the same lines of evidence, although they reached different conclusions from the evidence. Further, the defense experts generally agreed with the proposition that, in looking at the issue of causation, it is appropriate to consider all of the pertinent evidence and not just the epidemiological studies. The judge also criticized plaintiffs' experts for their skepticism about the use of meta-analysis. However, the Manual cautions that "when meta-analysis is applied to observational studies - either case-control or cohort - it becomes more controversial" due to the "methodological differences among studies." Reference Manual, supra, at 607.

In summary, the purpose of a Kemp hearing is to weed out "junk science," not to shield jurors from hearing expert testimony that is scientifically-based but unpersuasive to the trial judge. Landrigan, supra, 127 N.J. at 417; Kemp, supra, 174 N.J. at 427. "[R]egardless of a trial judge's view of the weight a party's

evidence deserves, the judge should trust the jury to evaluate witness credibility and decide what weight to give each side's evidence." State v. Stubblefield, __ N.J. Super. __, __ n.6 (App. Div. 2017) (slip op. at 21 n.16). It is the jury's core function to weigh the credibility of expert witnesses, and the trial court should not use a Kemp hearing as a vehicle to dismiss a case the court perceives as weak." Vigorous cross-examination, presentation of contrary evidence, and careful instruction on the burden of proof are the traditional and appropriate means of attacking shaky but admissible evidence." Daubert, supra, 509 U.S. at 596, 113 S. Ct. at 2798, 125 L. Ed. 2d at 484.

We conclude that the trial court misapplied its discretion in barring Dr. Kornbluth and Dr. Madigan from testifying. Accordingly, the orders entered in A-4698-14 and A-0910-16, barring their testimony and dismissing the complaints on summary judgment, are reversed and these cases are remanded to the trial court for further proceedings. We do not retain jurisdiction.

Reversed and remanded.

I hereby certify that the foregoing
is a true copy of the original on
file in my office.


CLERK OF THE APPELLATE DIVISION