United States Court of Appeals
For the First Circuit

No. 10-2301

UNITED STATES, ex rel. Kenneth James Jones,
Plaintiff, Appellant,

PRISCILLA PITT JONES and KENNETH JAMES JONES,
Plaintiffs,

v.

BRIGHAM AND WOMEN'S HOSPITAL, et al.,
Defendants, Appellees,

HARVARD MEDICAL SCHOOL, et al.,
Defendants.

APPEAL FROM THE UNITED STATES DISTRICT COURT
FOR THE DISTRICT OF MASSACHUSETTS

[Hon. William G. Young, U.S. District Judge]

Before

Lipez, Ripple*, and Howard,
Circuit Judges.

Jeremy L. Friedman, with whom William D. Hughes, Michael D. Kohn, Hughes & Nunn LLP, Kohn, Kohn & Colapinto, LLP, and Law Office of Jeremy L. Friedman, were on brief, for appellant.

Alan D. Rose, with whom Lisa A. Tenerowicz, and Rose, Chinitz & Rose, were on brief, for appellees.

May 7, 2012

* Of the Seventh Circuit, sitting by designation.

LIPEZ, Circuit Judge. This case requires us to address claims filed pursuant to the False Claims Act, 31 U.S.C. § 3729 (the "FCA"), alleging that the defendants submitted an application to the National Institute on Aging ("NIA") for research on Alzheimer's disease ("AD") which relied on falsified data. The district court granted summary judgment for the defendants. We vacate that ruling.

I.

In 2006, Dr. Kenneth Jones ("Jones" or "Relator") filed a qui tam action under the FCA against defendants Brigham and Women's Hospital ("BWH"), Massachusetts General Hospital ("MGH"), Dr. Marilyn Albert, and Dr. Ronald Killiany (collectively, the "Defendants"). Jones claimed that the Defendants violated the FCA by including false statements in a grant application that was submitted to the NIA, an institute within the National Institutes of Health ("NIH"). The NIH is an agency of the United States Department of Health and Human Services. Jones alleged that statements in the Program Project Grant Application (the "Application") were predicated on falsified data and that the Defendants, knowing of this falsity, failed to take corrective action or disavow the data. After the parties filed cross-motions for summary judgment, the district court granted judgment for the Defendants.

Jones timely appeals, maintaining that material factual disputes remain concerning the Defendants' conduct. He asserts that the record supports a conclusion that the Defendants violated the FCA by (1) "knowingly submitting an application for a grant to the [NIH] that was based on falsified and fraudulently manipulated study data and false statements of blinded, reliable methodologies," and (2) receiving NIH funds while knowingly in violation of regulations that require applicant institutions to investigate and report allegations of scientific misconduct.[1]

After careful review of the record, we conclude that the district court abused its discretion by excluding or failing to consider certain expert testimony. It then committed an error of law by failing to consider statements of the parties and experts in a manner required by the summary judgment standard. When properly considered, those statements generate genuine issues of material fact concerning some of Relator's FCA claims. We begin our explanation of these conclusions by describing the research project in question, Relator's concerns, and the NIH grant application process. We then recount in some detail the contents of the parties' expert reports. In Part II, we consider the district

[1] Harvard Medical School, Harvard University, and Dr. Marie F. Kijewski were also named as defendants in this action. The parties stipulated to voluntary dismissal without prejudice of the case against Harvard Medical School and Harvard University on March 17, 2009. Dr. Kijewski moved separately to dismiss the case against her. After a hearing on the matter, the district court granted Dr. Kijewski's motion to dismiss on July 10, 2009.

court's failure to examine Dr. Daniel Teitelbaum's expert report in its entirety, its improper evaluation of portions of other expert reports, and Relator's other claims.

A. The Research Project

The alleged false claims were submitted to the NIA in conjunction with a Program Project Grant ("PPG") proposal focused on AD, a neurodegenerative illness associated with aging. Dr. Marilyn Albert acted as the Principal Investigator (the "PI") on the PPG and oversaw the work of both Killiany and Jones.[2] The individual Defendants and Relator were part of a larger project team working to identify early physical manifestations of AD in certain regions of the brain and differentiate those characteristics from changes related to normal aging. Successfully doing so would enable health care providers to predict who will develop AD years before the individual displays diagnosable symptoms, thereby permitting early intervention. The NIH began funding this research in 1980 under a grant entitled "Age-related changes of cognition in health and disease" (the "Grant"), and continued to fund the project through 2007.

The proposed PPG consisted of four "Projects," long-term research studies focused on related issues, and four "Cores," each of which provided various types of support to the Projects.

[2] Both BWH and MGH were parties to the original grant. MGH has since merged with BWH. Albert and Killiany were both affiliated with MGH.

Killiany, a neuroanatomist, would head "Project 3," and utilize MRI to explore regions of interest ("ROIs"), including the EC, in the brain. Project 3 was a continuation of a study already in progress, the preliminary results of which were published in a 2000 paper authored by Killiany, Albert, and Dr. Mark Moss. During the study, Killiany and another researcher ("rater"), Dr. Teresa Gomez-Isla, developed and agreed to a protocol by which to locate and outline the EC and other ROIs on MRI scans. According to the Application, raters were blinded, meaning they were unaware of participants' cognitive groupings.[3] Raters were provided with participants' identification numbers and sometimes their names. Participant identification numbers were assigned consecutively upon participants' entry into the study; thus the numbers did not communicate information to the raters about participants' cognitive statuses. Using a trackball mouse and "Neuroview" software, the raters manually outlined a number of brain structures, including the EC, on 103 participants' MRI scans.[4] Based on the raters'

[3] Study participants were assigned to one of three diagnostic groups - normal, questionable, or converter. Initially, all participants were categorized as either "normal" (also known as "control") or "questionable." Participants were assigned to the normal group if they met the clinical dementia rating for normal cognition or were assigned to the questionable group if they met the clinical dementia rating for questionable cognition. Over the course of the study, some participants' cognitive difficulties progressed such that they met the clinical criteria for probable AD. These participants were reassigned to the "converter" group.

[4] Between approximately 1995 and 1999, Killiany traced the scans of 103 participants. In approximately 1996 or 1997, Gomez-

outlines, the software calculated the volume of the traced EC, and those calculations were later sent to Dr. Mary Hyde, the Data Manager and Programmer for "Core B," the Data Management and Statistical Core.[5] Thereafter, members of Core B, supervised by Jones, conducted statistical analyses of the data to determine whether changes in the volume of the EC could help predict which cognitively healthy people would develop AD in the future.

Between 1995 and 1999, Killiany modified his outlining process. In his deposition, Killiany testified that as he worked through participants' scans, he encountered a number of "anatomical anomalies." When he encountered such anomalies, he went back and reviewed earlier outlines to ensure that those tracings properly considered the anatomical issue. If, upon reviewing an outline, he felt that it should be revised, he re-traced the EC boundary, the software recalculated the volumetric data, and he eventually sent the new data to Hyde.

Isla traced the scans of a subset of 25 participants. In 1996 or 1997, members of Core B compared Gomez-Isla's tracings to Killiany's tracings of the same 25 participants to determine inter-rater reliability, an index of how close the volumetric data generated by one rater's measurements is to that generated by another rater's measurements of the same item.

[5] Volumetric data was sent to Core B for statistical analysis periodically over the course of several years. Killiany stated that he intermittently sent data to Hyde in batches, which sometimes included re-measurements. Gomez-Isla stated that she did not recall personally sending data to Core B.

B. Relator's Concerns About the Study Data

As noted, the relator in this case, Jones, headed Core B. In that role, he supervised data management, assessed project progress, analyzed project data, and developed new analytic frameworks. In March 2001, Jones met with Albert and Dr. Keith Johnson, the leader of Project 2. Jones and Johnson alerted Albert that they had concerns about the data that Killiany produced prior to 1998, which had been used to demonstrate a statistically significant relationship between the volume of the EC and conversion to AD. Jones noted that there was more than one data set for a number of study participants and expressed concern regarding the quantitative differences between those sets. Killiany had re-measured the scans of 30 participants - thirteen in the normal group, twelve in the questionable group, and five in the converter group.[6] Jones had particular concerns about 23 of those re-measurements. Jones told Albert that the alterations were substantial and that, according to his analysis, the alterations were responsible for the apparent statistical significance. Jones asked Albert to look into the issue.[7]

---

[6] There is a dispute about the precise number of scans Killiany re-measured and their categorization. Defendants claim that 31 scans were re-traced, of which twelve were in the normal group, eleven were in the questionable group, two were in the converter group, and six were unclassified. On summary judgment, we resolve such disputes in favor of the non-moving party.

[7] The parties disagree as to whether Jones asked Albert to have someone "verify" the data, as appellees argue, or "re-measure"

In response, Albert asked Dr. Mark Moss, a noted neuroanatomist who had been involved with the PPG since its inception, to examine for accuracy the 23 re-measurements about which Jones and Johnson expressed concern. Moss reviewed the scans in question and hand wrote notes expressing his opinion of the accuracy of each scan. He concluded that with one exception, Killiany's second set of measurements more accurately outlined the EC. At Albert's request, Moss gave his notes to Killiany. Later, Killiany created a typewritten document ostensibly containing Moss's notes,[8] and sent the document to Jones and Johnson. Albert also received the typed notes and was satisfied with Moss's conclusion that the re-measurements were more accurate than the initial ones. Jones was unsatisfied with Moss's assessment and asked that the scans in question be re-measured. Albert refused. In her deposition, Albert stated that she believed that an

---

the EC on the scans, as appellant argues. In this case, we do not believe that the distinction is meaningful. Whether Albert was asked to have the data verified or the scans re-measured, she was on notice that Jones and Johnson had concerns about Killiany's second set of data and, therefore, had reason to look into the matter.

[8] Since Moss's handwritten notes no longer exist, it is impossible to determine whether Killiany's typed notes are an exact replica. In his deposition, Moss stated that he likely would not have used certain phrases that appear in the version that Killiany typed. Moss said that due to the amount of time that had passed, he could not be sure that the language that appeared in the typed version was what he said in his handwritten notes, but he also noted that he trusted Killiany and "assume[d]" that the typed notes were accurate.

additional set of measurements would confuse instead of answer the question of which data set was more accurate, so she did not have the scans re-measured. Jones continued to work with data generated from Killiany's tracings until July 30, 2001, when he told Albert that he would no longer work with Killiany's data without "neutralizing" it. After that, Jones did not work with any MRI data generated from Killiany's tracings.

C. The NIH Grant Application Process

To secure funding from the NIH for age-related research, institutions must submit applications to the Center for Scientific Review and the NIH. The applications are then forwarded to the NIA, where they first undergo a peer review process conducted by a panel of independent experts in the relevant field. The panel considers a number of factors, including the quality and originality of the science proposed, the quality of the investigators, and the quality of the facilities in which the research will take place. Based on these considerations, the Scientific Review Administrator prepares a statement (the "Pink Sheets") summarizing the strengths and weaknesses of a proposal. If the project is recommended for further consideration, the National Advisory Council on Aging will review the application, focusing its evaluation on the perceived scientific quality of the application. After the Advisory Council's review, final approval is committed to the discretion of the Director of the NIA.

On October 1, 2001, Albert and MGH submitted a PPG Application for the 2002-2007 NIH funding cycle. As part of the Application, Defendants described the results of relevant preliminary studies, including the MRI study in which Killiany allegedly manipulated data. The Application also outlined the methods that would be used in future research and the protocols that the researchers planned to employ to ensure data reliability. The Defendants did not include the allegedly false underlying data itself, but did include a discussion of the results generated by the data and explained why those results supported the proposed study. The Application did not mention the existence of two sets of data or Jones's allegations of wrongdoing. Before submission, Albert and a representative of MGH certified the truthfulness of the Application's contents on its cover.[9]

---

[9] Albert signed the following certification:

> I certify that the statements herein are true, complete and accurate to the best of my knowledge. I am aware that any false, fictitious, or fraudulent statements or claims may subject me to criminal, civil, or administrative penalties. I agree to accept responsibility for the scientific conduct of the project and to provide the required progress reports if a grant is awarded as a result of this application.

MGH representative Marcia L. Smith signed the following certification:

> I certify that the statements herein are true, complete and accurate to the best of my knowledge, and accept the obligation to comply with Public Health Service terms and conditions if a grant is awarded as a result of this application. I am aware that any false, fictitious, or fraudulent statements or claims may subject me to

D. The Parties' Experts

1. Defense Expert Dr. Andrew J. Saykin

Saykin is a professor of radiology at the Indiana University School of Medicine and Director of the Indiana University Center for Neuroimaging, with a research focus on "the integration of neuroimaging and genomic data with emphasis on early detection of Alzheimer's disease."[10] In his report, Saykin stated that outlining "ROIs on brain scans inevitably involves a learning curve for the neuroanatomic rater and some degree of trial and error, especially as methods are refined and solidified." As such, Saykin did "not believe that it was unusual or inappropriate for Killiany to perform re-measurements to improve his accuracy of measurement, as long as he remained blinded to the clinical status of the participants." Saykin stated that re-measurements like those that Killiany made are normal and need not be reported "unless the re-measurements were undertaken as part of a formal reliability assessment after a change in the protocol for measuring the EC." He also stated that he believed that "Killiany took great pains to refine and update his measurements, following the protocol or procedural guidelines he had developed, in an attempt to make

---

criminal, civil, or administrative penalties.

[10] Saykin held these positions as of August 4, 2010, the date of his expert report.

the measurements more accurate." Saykin concluded that Killiany was blinded at the time of the re-measurements.

Regarding the allegations of misconduct, Saykin stated that it was "appropriate for [Albert] to ask a prominent neuroanatomist to review Killiany's measurements and provide his analysis demonstrating whether the re-measurements were done in an attempt at greater accuracy or simply to prove the hypothesis. Since Albert concluded the former had occurred . . . there was no need to report what had happened to anyone." Saykin stated that if, in fact, "Killiany's first set of measurements did not demonstrate the predictive value of the volume of the EC, those measurements were inconsistent with the findings from many other scientific studies," as "[i]t is now scientifically accepted that the volume of the EC is a good predictor of conversion to AD."

2. Relator Expert Dr. Norbert Schuff

Not surprisingly, Relator's expert offered a contrasting conclusion. Schuff is a professor of radiology at the University of California in San Francisco, an investigator at the VA Medical Center in San Francisco, lead physicist at the Center for Imaging of Neurodegenerative Diseases at the VA Medical Center in San Francisco, and a researcher focusing on the development of new MRI methods and concepts to identify markers of neurodegenerative

diseases, including AD.[11] He stated that while Killiany and the other rater, Gomez-Isla, had reached an "initial conceptual understanding [of] how to trace the [EC]" and had achieved "high inter-rater reliability" using that protocol for the first set of measurements, Killiany's re-measurements included "extensive" revisions that were "frequently inconsistent with the initially adapted protocol . . . and seemed to introduce greater rater variability." Schuff noted that any changes to the protocol should have been discussed with Gomez-Isla, after which Gomez-Isla would have revised her tracings and the statistics Core would have retested reliability. Because Killiany failed to take these steps, Schuff concluded that the only "explanation for the highly selective revisions of the [EC] tracings in [the normal group]" was that such "revisions were made with knowledge of the [participants'] diagnos[e]s." Schuff noted that the objectivity of the EC measurements was material to the NIH's and the peer reviewers' assessments of the merit of the proposed project.

3. Relator Expert Dr. Martha Isabel Dávila-García

Dávila-García, an Associate Professor at Howard University College of Medicine,[12] explained the NIH application consideration process based on her knowledge as a past application

[11] Schuff held these positions as of the date of his expert report, August 2, 2010.

[12] Dávila-García held this position as of the date of her expert report, September 15, 2010.

reviewer. She stated, among other things, that the Application contained "a number of statements that were material to the government's funding decision . . . includ[ing] the preliminary data and progress report of the scientists' ongoing research project, as well as the reliability of the blinded methodologies the scientists claimed they followed in validating that data." Dávila-García further suggested that "[e]ach of these statements was required to be made in the grant application, and each is fundamental to the peer review ranking of the application."

Dávila-García also stated that Albert's inquiry into the alleged misconduct was insufficient. Under 42 C.F.R. § 50.103(d)(8)-(9) (2001),[13] Albert was required to seek out someone free from "real or apparent conflicts of interest" with "appropriate expertise to carry out a thorough and authoritative evaluation of the relevant evidence in any inquiry or investigation." Dávila-García opined that Moss was an improper

[13] Federal regulation 42 C.F.R § 50.103 describes the responsibilities of awardee and applicant institutions that seek or receive assistance from the Public Health Service division of the United States Department of Health and Human Services, of which the NIH is an agency. Subsection 50.103(d) describes the various assurances that subject institutions must make to ensure that they are capable of dealing with and reporting possible scientific misconduct, including "[s]ecuring necessary and appropriate expertise to carry out a thorough and authoritative evaluation of the relevant evidence in any inquiry or investigation," § 50.103(d)(8), and "[t]aking precautions against real or apparent conflicts of interest on the part of those involved in the inquiry or investigation," § 50.103(d)(9). This regulation has since been replaced by 42 C.F.R. § 93.100-319.

choice to review Killiany's work, as Moss "was Killiany's boss and mentor, and the chairman of the Department of [A]natomy and [N]eurobiology at Boston University School of Medicine," who "would appear to . . . have a conflict of interest and be biased to protect the reputation of his Department and its members."[14]

4. Relator Expert Dr. Daniel Teitelbaum

Teitelbaum has a Ph.D. in engineering and has experience working with "statistics, data analysis, predictive modeling, building computer simulation, mathematical optimization, [and] logistics and operations research." He stated that in his opinion, "the altered data points undo the validity of the study's conclusions" because "[r]ather than being a systematic visitation upon the study data, Killiany conducted the second set of measurements by cherry-picking the study subjects in a non-random fashion." Moreover, Teitelbaum stated that "the changes themselves were responsible for the significance of the results Killiany claimed to achieve . . . . Had the original data been used, Killiany could not have reported his findings in published

[14] Killiany has held a number of positions in the Department of Anatomy and Neurobiology at Boston University School of Medicine. He was a postdoctoral fellow there from 1991-92, a research associate from 1994-2001, a research assistant professor from 2001-2003, an assistant professor from 2003-2006, and, as of 2006, an associate professor in the department. According to Killiany's curriculum vitae, he has worked on at least four grants for which Moss was the principal investigator, and they have authored more than thirty book chapters and peer-reviewed publications together.

scientific journals or to the NIH in support of an application for a Program Project Grant."

Regarding the reliability study, Teitelbaum explained that

> [w]hile minor refinements in measures may be appropriate after a reliability study is performed, it would be inappropriate to make significant alterations or to record data in deviation of the protocols subjected to the reliability study. It would be particularly unusual for there to be any changes to statistical measurements without a full, documented explanation as to why such changes were required. In this case, the magnitude of the changes clearly conflict [sic] with the extent of normal variations based on a reported 0.96 Pearson Correlation achieved.[15]

Teitelbaum concluded "that the data in Killiany's study which was reported and relied upon in the 2002 PPG Renewal Application was deliberately manipulated in order to achieve statistically significant results that could not be achieved but for the manipulation of the data." He stated that as a result, "Killiany reported data that had not truthfully been generated pursuant to the stated protocols, and he negated his assertion that those protocols had been proven to be blinded and reliable."

[15] The Pearson Correlation coefficient is a statistical value between -1.0 and 1.0 used to indicate the linear relationship between two variables. A coefficient of 0.0 indicates that there is no relationship between the variables, and a coefficient of 1.0 indicates a perfect positive linear relationship between the variables.

II.

When Jones filed his claim in 2006,[16] the FCA imposed civil liability on any person who either "knowingly presents, or causes to be presented to an officer or employee of the United States Government . . . a false or fraudulent claim for payment or approval," 31 U.S.C. § 3729(a)(1), or "knowingly makes, uses, or causes to be made or used, a false record or statement to get a false or fraudulent claim paid or approved by the Government," id. § 3729(a)(2).[17] A person acts "knowingly" if he or she "(1) had actual knowledge of the information; (2) acts in deliberate ignorance of the truth or falsity of the information; or (3) acts in reckless disregard of the truth or falsity of the information."

---

[16] Under the FCA, the Attorney General may bring an action, 31 U.S.C. § 3730(a), or an individual may bring an action "in the name of the Government," 31 U.S.C. § 3730(b)(1). An individual who brings an action under the FCA, who is known as a relator, may proceed regardless of whether the Government chooses to intervene and take over the action. 31 U.S.C. § 3730(b)(4)(B).

[17] Our discussion relies on the provision in effect at the time Jones filed his complaint. See United States ex rel. Susan Hutcheson and Phillip Brown v. Blackstone Medical, Inc., 647 F.3d 377, 380 (1st Cir. 2011), cert. denied, 132 S. Ct. 815 (2011). The FCA has since been amended by the Fraud Enforcement Recovery Act ("FERA") of 2009, Pub. L. No. 111-21, 123 Stat. 1617 (2009). Under FERA, an individual incurs liability when he or she "knowingly presents, or causes to be presented, a false or fraudulent claim for payment or approval," 31 U.S.C. § 3729(a)(1)(A), or "knowingly makes, uses, or causes to be made or used, a false record or statement material to a false or fraudulent claim," id. § 3729(a)(1)(B). Application of the amended language would not affect the outcome in this case.

Id. § 3729(b).[18] A relator need not show any "proof of specific intent to defraud." Id. We have also "long held that the FCA is subject to a judicially-imposed requirement that the allegedly false claim or statement be material." United States ex rel. Loughren v. Unum Grp., 613 F.3d 300, 307 (1st Cir. 2010).

The district court found that Jones had not generated genuine issues of material fact on any of the elements at issue - falsity, materiality, and knowledge. Jones argues to the contrary. He notes that the Application relies on Killiany's research and claims that Killiany "fraudulently altered the MRI study data prior to 1998 to produce false results of a statistically significant correlation between conversion to AD and volume of the EC," and did so "after the scientists had conducted a reliability study showing consistency in the manual drawings of EC boundaries." Jones claims that the alterations led to results that were incompatible with the results of the reliability study that was conducted and with the stated protocols. He alleges that "Albert and the defendant

[18] The parties do not distinguish how each theory of liability applies to each particular defendant; we leave that issue open for remand. We note, though, that while Killiany and Albert may have been directly responsible for development of the alleged falsities at issue, MGH and BWH, which has since acquired MGH, may be vicariously liable under the FCA for the misrepresentations of their employees. We have long held that corporate defendants may be subject to FCA liability when the alleged misrepresentations are made while the employee is acting within the scope of his or her employment. See United States v. O'Connell, 890 F.2d 563, 568 (1st Cir. 1989) (noting that nothing in the FCA's text proscribes vicarious liability, and application of vicarious liability serves the FCA's dual purposes of restitution and deterrence).

hospitals submitted the subject application for NIH funding on the entire PPG, including a separate MRI project, without disclosing to NIH the second set of MRI data, its statistical impact or the allegations and appearance of Killiany's misconduct."[19] Jones also asserts a series of separate errors that he claims undermine the court's ultimate conclusion.

We review the district court's evidentiary determinations, namely, its decisions to admit or exclude expert testimony, for abuse of discretion. Alt. Sys. Concepts, Inc. v. Synopsys, Inc., 374 F.3d 23, 31 (1st Cir. 2004); see also Gen. Elec. Co. v. Joiner, 522 U.S. 136, 142-43 (1997) (noting that abuse

---

[19] Jones also claims that the district court abused its discretion by not finding that the Defendants had spoliated evidence, including scientific notebooks, MRI images, data files, emails, and any record of the internal inquiry into the allegations of scientific misconduct. Jones argues that such evidence should have been preserved under 45 C.F.R. § 74.53(b), "Post-Award Requirements Reports and Records." Under § 74.53(b),

> [f]inancial records, supporting documents, statistical records, and all other records pertinent to an award shall be retained for a period of three years from the date of submission of the final expenditure report . . . .

The district court found - as the name of the regulation suggests - that the Post-Award Requirements are forward-looking and apply only once a grant has been awarded. Thus, because the grant in question was not funded until 2002, the Post-Award Requirements did not apply to records pre-dating 2002. We agree with the district court that Jones's spoliation claim fails on this basis. The district court also noted that applicant institutions are required to retain documents related to misconduct inquiries for three years, a time span that terminated in 2004 in this case. Jones gives us no basis to doubt this determination; indeed, we note that Jones first brought suit in 2006.

of discretion review applies to threshold evidentiary determination made in connection with summary judgment motions). "Evidentiary rulings have the potential to shape and winnow the scope of the summary judgment inquiry, and a trial court should have as much leeway in dealing with those matters at the summary judgment stage as at trial." Alt. Sys. Concepts, Inc., 374 F.3d at 31-32. A court abuses its discretion if it commits "a material error of law," or if it "ignores a material factor deserving significant weight, relies upon an improper factor, or assesses only the proper mix of factors but makes a serious mistake in evaluating them." Downey v. Bob's Disc. Furniture Holdings, 633 F.3d 1, 5 (1st Cir. 2011).

After reviewing the district court's evidentiary determinations and thereby settling the scope of the summary judgment record, we review the court's grant of summary judgment de novo. Schubert v. Nissan Motor Corp. in U.S.A., 148 F.3d 25, 29 (1st Cir. 1998); see also Sch. Union No. 37 v. United Nat'l Ins. Co., 617 F.3d 554, 558 (1st Cir. 2010). "[W]e will reverse a grant of summary judgment only if, making all factual inferences in favor of the non-moving party, a rational fact-finder could resolve the legal issue for either side." D&H Therapy Assocs., LLC v. Boston Mut. Life Ins. Co., 640 F.3d 27, 34 (1st Cir. 2011). Where, as here, the parties have filed cross-motions for summary judgment, we must "determine whether either of the parties deserves judgment as

a matter of law on facts that are not disputed." Sch. Union No. 37, 617 F.3d at 559 (quoting Littlefield v. Acadia Ins. Co., 392 F.3d 1, 6 (1st Cir. 2004)) (internal quotation mark omitted).

A. Teitelbaum's Expert Testimony

The Defendants filed a motion in limine to preclude Relator from offering certain testimony and evidence, including aspects of the proposed testimony from each of his three expert witness reports. Defendants argued, among other things, that as a statistician, Teitelbaum was unqualified to assess inter-rater reliability or to opine when a reliability study should be conducted or what would be expected on the basis of reliability results. The Defendants also challenged the admissibility of various statements in Teitelbaum's report, arguing that his opinions lacked sufficient support or were ambiguous, misleading, or otherwise inadmissible.

Jones notes that the district court did not directly address the motion in limine in the memorandum accompanying its summary judgment ruling, and, indeed, did not mention Teitelbaum's report or its admissibility at all. Although a district court is afforded great discretion in deciding whether to admit or exclude opinion evidence, Crowe v. Marchand, 506 F.3d 13, 16 (1st Cir. 2007), it cannot abdicate that responsibility altogether. In this case, the district court could not properly conduct its summary judgment analysis without determining the admissibility of

Teitelbaum's report, which speaks directly to issues at the heart of Jones's claims. See, e.g., Cruz-Vázquez v. Mennonite Gen. Hosp., 613 F.3d 54, 57 (1st Cir. 2010) (noting that it is the district court's responsibility to "determin[e] whether to admit or exclude expert testimony" based on an evaluation of whether "the expert's testimony both rests on a reliable foundation and is relevant to the task at hand" (internal quotation marks omitted)). By failing to exercise its discretion, namely, failing to admit or exclude Teitelbaum's report, the district court committed an error of law and, thereby, abused its discretion. See Downey, 633 F.3d at 5.

That error, however, does not necessarily mean that we must vacate the district court's summary judgment ruling. In the absence of the district court's analysis regarding the dispute over Teitelbaum's qualifications, we will make our own determination on the admissibility of Teitelbaum's testimony in order to determine the scope of the summary judgment record. See Boston Duck Tours, LP v. Super Duck Tours, LLC, 531 F.3d 1, 15 (1st Cir. 2008) (stating that an appellate court may make a determination on "a relevant and required issue" where remanding "would be a waste of judicial resources and incompatible with the urgency of the issue before us").

According to the curriculum vitae submitted with his expert report, Teitelbaum has a Ph.D. in Engineering and Public

Policy. Since his graduation in 1998, he has worked in a variety of settings with duties related to statistics, data analysis, predictive modeling, computer simulation, mathematical optimization, and logistics and operations research. In rendering his opinions, Teitelbaum stated that he analyzed a variety of materials including, among other things, Killiany's original and revised data sets, Killiany's 2000 paper, Albert's deposition testimony, Relator's Table, a data chart illustrating the revisions' effect on reliability, and excerpts from the Application.

We see no bar to the admission of Teitelbaum's testimony in light of this dispute over his qualifications. In our judgment, that dispute only goes to the weight of his opinion testimony before a fact-finder. For purposes of summary judgment, having determined that Teitelbaum's report is admissible, we ask whether his report, along with other record evidence, generates genuine issues of material fact as to whether Killiany falsified data while unblinded to participants' group membership, and, thereby, undermined the study's results and statements made by the Defendants in the PPG Application.

B. Falsity

Jones alleged that Defendants made three different misrepresentations in conjunction with the NIH Application. First, Jones claimed that the Defendants described and relied on

Killiany's research without reflecting the alleged fraudulent manipulation of the EC tracings, specifically, the unblinded, selective enlargement of certain tracings to produce results that exaggerated group differences.

Second, Jones alleged that the Defendants made false statements regarding reliability methodologies in the Application by reporting a 0.96 Pearson Correlation coefficient. Jones argued that this coefficient corresponded to the first set of tracings that Killiany made, even though the second set was the data that produced the statistically significant result that the Defendants relied upon in the Application. Jones argued that when a reliability study was conducted comparing Gomez-Isla's tracings and Killiany's second set of measurements, the correlation coefficient dropped to 0.54.

Finally, Jones alleged that the Defendants violated the FCA by falsely certifying that they were in compliance with Public Health Services ("PHS") terms and conditions and NIH Scientific Misconduct Regulations. Jones alleged that Defendants failed to meet their obligations under 42 C.F.R. § 50.103(c)(3), which requires applicant institutions to take "immediate and appropriate action as soon as misconduct on the part of employees or persons within the organization's control is suspected or alleged." Jones claimed that Albert's inquiry into Jones's allegations was insufficient and that the Defendants were obligated to conduct a

full investigation and report the results of that investigation to the NIH.

The district court considered the parties' claims under an FCA falsity framework that we have since rejected. The court stated:

> There are three theories under which a claim may be "false or fraudulent" under the Act. These are: (1) factual falsity; (2) legal falsity under an express certification theory; and (3) legal falsity under an implied certification theory.

(Footnotes omitted.) After the parties briefed this appeal, we had occasion to clarify the proper framework for analyzing FCA claims. See generally Hutcheson, 647 F.3d 377; see also New York v. Amgen, Inc., 652 F.3d 103 (1st Cir. 2011), cert. dismissed, 132 S. Ct. 993 (2011). In Hutcheson, we rejected rigid divisions between factual and legal falsity, and express and implied certification, noting that the text of the FCA does not make such distinctions. The use of such categories, in "our view[,] . . . may do more to obscure than clarify the issues before us." Hutcheson, 647 F.3d at 385-86. Instead, we take a broad view of what may constitute a false or fraudulent statement to avoid "foreclos[ing] FCA liability in situations that Congress intended to fall within the Act's scope." Id. at 387 (quoting United States v. Sci. Applications Int'l Corp., 626 F.3d 1257, 1268 (D.C. Cir. 2010)) (internal quotation marks omitted). Taking this broad view does not, however, create limitless liability. Indeed, FCA liability continues to be

circumscribed by "strict enforcement of the Act's materiality and scienter requirements." Id. at 387-88 (quoting Sci. Applications Int'l Corp., 626 F.3d at 1280) (internal quotation marks omitted). We analyze each of Relator's claims by applying the plain language of the FCA in accordance with our decision in Hutcheson.

1. Falsified Data Generated by Unblinded Methodology

a. Data Generated from Re-measurements

The district court found that Jones "failed to articulate how the supposedly false data relates to a false statement in the Application," as "[t]here is no evidence that the EC data itself was submitted as part of the Application." Moreover, the district court found that "the act of tracing the boundaries of the EC is subjective and requires the exercise of scientific judgment" such that "two scientists who use the same protocol manually to trace the EC may nevertheless obtain different results." In dismissing the significance of the disagreement of the experts, the district court noted that such disputes over the exercise of scientific judgment may not form the proper basis for an FCA claim and cannot "yield a resolution where one can state with reasonable certainty that one conclusion is true and the other false."

Although it is true that the allegedly false EC volumetric data was not itself included in the Application, that fact is not determinative of the false claim allegation. The statute makes it a violation to "use[] . . . a false record or

statement to get a false or fraudulent claim paid or approved by the Government." 31 U.S.C. § 3729(a)(2). A number of statements in the Application demonstrate reliance on the study's conclusions and therefore necessarily implicate the allegedly false data. For example, the Application contains the following statements:

> [W]e found that 3 measures obtained at baseline were highly significant predictors of who would develop AD on follow-up: (1) the volume of the entorhinal cortex, (2) the banks of the superior temporal sulcus, and (3) the caudal portion of the anterior cingulate. When the control subjects were compared with the non-demented individuals with memory impairments who ultimately developed AD (i.e., the 'converters'), the accuracy of discrimination was 93% based on the MRI measure at baseline (sensitivity = 0.95; specificity = 0.90). . . . One hundred percent (100%) of normal controls could be discriminated from the patients with mild AD based on these 3 baseline MRI measures (Killiany et al., 2000).
>
> [W]e have demonstrated that to identify individuals in the prodromal phase of AD, regions such as the entorhinal cortex and the caudal portion of the anterior cingulate are highly discriminating.
>
> Our major finding is that measures of memory and executive function, or SPECT and MRI measures of brain regions related to these domains (such as the entorhinal cortex, the hippocampus, and the caudal portion of the anterior cingulate) are highly predictive of subsequent development of dementia among non-demented individuals with memory problems. . . . In the current application we propose to expand the longitudinal evaluation of the subjects to permit a more detailed understanding of the variables that predict course during prodromal AD.

> The most discriminating MRI measures pertain to atrophy of the medial temporal lobe (particularly the entorhinal cortex), and the volume of anterior and posterior cingulate (Killiany et al., 2000).

These statements rely on the data challenged by Jones as false. In the language of the FCA, they "use . . . a false record." Thus premised, the statements would not be "true, complete and accurate" as required by the certifications signed by Albert and MGH.

We agree with the district court that "[e]xpressions of opinion, scientific judgments, or statements as to conclusions about which reasonable minds may differ cannot be false." (citing United States ex rel. Roby v. Boeing Co., 100 F. Supp. 2d 619, 625 (S.D. Ohio 2000)). However, we disagree that the creation of the data in question was necessarily a matter of scientific judgment. The district court relied on "the undisputed fact that tracing the EC is highly subjective and thus two scientists who use the same protocol manually to trace the EC may nevertheless obtain different results." This reliance, however, misses the point that the various results produced in this case were obtained by one scientist purportedly using the same protocol. Although the decision as to which measurement method to employ was a question of scientific judgment, that is not the issue here. As Schuff noted, Killiany and Gomez-Isla "had reached a conceptual understanding [about] how to trace the [EC]" and used that protocol in their initial measurements, which demonstrated high inter-rater

reliability. The real issue is whether, as Schuff opines, Killiany's revisions "substantially deviate[d] . . . from the initial protocol for [the EC] that [he and Gomez-Isla] had established to the point that the initial and new markings [were] no longer consistent" or capable of meaningful comparison. Indeed, Killiany himself explained that one aim of the project was to substantiate whether "two knowledgeable individuals" could "apply a definition of the [EC] . . . across a large number of MRI scans" and "actually even agree on where the [EC] would be."

Killiany suggested that over the course of measuring 103 participants' brain scans, he went through a learning curve during which he discovered various anatomical anomalies that required modifications of his measurement technique. He explained that after he came across this type of anomaly, he would review previously outlined scans and evaluate them in accordance with his now-modified technique. Although the Defendants' expert Saykin stated that such re-measurements to improve accuracy were not "unusual or inappropriate . . . as long as [Killiany] remained blinded to the clinical status of the participants," the record raises questions about Killiany's explanation.

The distribution of altered data among and within participant groups raises the greatest concern. Of the 103 participants in the final study population, 30 participants' scans were re-traced. Teitelbaum suggested that if the changes were in

fact revisions for accuracy, one would expect some re-tracings to be smaller than the corresponding original tracing, resulting in a lower volume measurement, and some re-tracings to be larger, resulting in a higher volume measurement. Moreover, Teitelbaum stated, one would expect to see such enlargements and reductions occurring randomly in all groups of participants. Instead, the most frequent and dramatic changes occur in the normal group. Measurements were changed for 13 of 24 normals (54.2 percent), with volume increases from 0.3 to 283.3 percent.[20] Of the 19 "converters," five were changed (26.3 percent), with volume increases between 0.3 and 12.1 percent.[21] Of the 60 participants classified as "questionables," twelve were changed (20 percent), with volume changes between -23.9 percent and 72.5 percent.[22] Moreover, the changes within the normal group did not appear random. The six smallest initial measurements were each increased by over 100 percent. Twelve of the smallest thirteen original measurements were revised, while only one of the eleven largest was (and that one by only 0.3 percent). Of course, it may be that one of the anatomic anomalies that Killiany discovered affected only

[20] Participants categorized as "normals" or "controls" met the clinical dementia rating for normal cognition.

[21] Participants categorized as "converters" met the clinical criteria for probable AD.

[22] Participants categorized as "questionables" met the clinical dementia rating for questionable cognition.

the smallest of the original measurements. Teitelbaum suggested that if this were the case, however, he would expect to see revisions concentrated within the smallest original measurements in the other groups as well. Instead, changed data points in the converter and questionable groups appear throughout the range of initial measurements. Five changes in the questionable group occur in the smallest 1/3 of measurements; four in the middle 1/3; and three in the largest 1/3. In the converter group, only one change appears in the nine smallest measurements. The remaining four changes occur in the largest ten measurements.

In essence, moderate revisions occurred seemingly randomly in the converter and questionable groups, but relatively large revisions were concentrated in the smallest half of the control group. As Schuff stated, most of Killiany's revisions were "quite extensive," "biased toward normal subjects," frequently "inconsistent with the initially adapted protocol agreed to by the raters, and "not founded on scientific reason." Schuff further explained that "[i]f Killiany had made the second set of measurements as part of his 'learning curve,'" sound scientific practice required him "to generate documentation and work papers in connection with his corrections . . . and share[] what he learned with his colleagues." Killiany testified that he does not recall discussing his decision to perform re-measurements with anyone on the PPG, and that after he initially discussed the EC boundaries

with Gomez-Isla, the two did not have subsequent discussion on the topic. The upward revisions in the control group were critical to the predictive value of the study; if normal subjects showed large EC volume changes and moved into the converter or questionable group, the result was physical evidence of potentially great predictive significance.

The revisions in question do not implicate questions of scientific judgment as the district court suggested, because all the measurements in question were purportedly generated by a single protocol that Killiany and Gomez-Isla agreed to before beginning the measurements. Indeed, as Schuff noted, whether Killiany's measurements were more or less accurate than the initial measurements is not at issue. Even if Killiany's re-measurements fall within an accepted range of scientific accuracy, a question remains as to whether the data was falsified by intentionally exaggerating the EC boundaries of normal subjects to achieve a desired result. We conclude that the distribution of revisions presents a genuine issue of material fact as to whether, as Teitelbaum put it, Killiany cherry-picked measurements to revise in a non-random fashion "in order to produce data that would support his hypothesis on the role of EC volume and the prediction of prodromal AD."

b. Blinded Methodologies

Using data provided by the defendants during discovery, Jones created a data table illustrating the distribution of alterations among and within participant groups ("Relator's Table" or "the Table"). The Table visually demonstrated the distribution and quantitative nature of the re-measurements. It listed the original and revised data for the 103 participants reported in Killiany's 2000 paper. On the left side of the Table, participants were ordered from smallest to largest according to their original measurements; on the right side of the Table, participants were ordered from smallest to largest according to their revised measurements. The significant volume enlargement of eleven normal participants is illustrated by arrows pointing from the participants' original measurements to their revised measurements. Most of the arrows start from the smallest original measurements and move downward to the same participants' revised numbers, which tend to be among the largest revised measurements. Relator claimed that the Table demonstrated that "Killiany selectively cherry-picked and substantially altered the normal subjects with the smallest [ECs]." Such selective alterations would necessarily indicate that Killiany was unblinded and aware of participants' group affiliations.

In his deposition, Killiany stated that he was blinded to the group status of study participants when he was making his

tracings and transmitting the volumetric data generated from those tracings. Albert stated that she believed that Killiany had stayed blinded because he did not have access to participants' diagnoses and because he told her that he had been blinded. Johnson stated that it was "[his] understanding . . . that the operator who is implementing the protocol . . . would be blinded to the classification of the subject being [traced]." Saykin concluded that "[b]ased on all the information [he] reviewed, [he] believed that Dr. Killiany did remain blind to the clinical status of the cases he was analyzing or re-analyzing anatomically, as would be standard and appropriate in this type of research."

Relying on Albert and Johnson's depositions,[23] the district court found that "[a]mple record evidence shows that Killiany was . . . blinded to the group status of the participants for which he traced the boundaries of the EC" and to "the statistical significance of any data he produced." The district

[23] Johnson, the leader of Project 2, along with Jones, first reported concerns about the data to Albert. On appeal, Defendants maintain that "Johnson confirmed that Killiany was blinded." However, the Johnson deposition excerpt that Defendants cite - and upon which the district court relied - confirms only that Johnson "was generally aware that those types of [blinding] procedures would be involved." When Johnson was asked who would know if blinding procedures had been violated, Johnson stated that he "would assume those involved with executing the protocol and those who would be overseeing that execution, which would be Dr. Killiany and perhaps Dr. Moss." It is apparent from Johnson's statement that he would not expect to hear if Killiany had breached the protocol. Thus, the fact that Johnson did not hear of such a breach does not conclusively resolve the question of whether Killiany was appropriately blinded.

court also noted that "Relator himself admitted at his deposition that he has no evidence that Dr. Killiany was not following proper, blinded methodologies when retracing EC boundaries."

On appeal, Jones maintains that he referred only to his direct personal knowledge when he indicated that he had no evidence that Killiany was not blinded. Moreover, Jones insists that his personal knowledge is not determinative. Rather, Jones argues, Teitelbaum's independent analysis of the raw data and the Relator's Table each created a genuine issue of fact as to whether Killiany was blinded.[24] We agree. In light of Teitelbaum's independent analysis and Relator's Table, there is a genuine issue of fact as to whether Killiany was blinded or instead deliberately selected certain participants whose data he manipulated to create the statistically significant result that he ultimately reached. The district court erred in concluding otherwise.

2. Reliability Study

Jones alleged that the Application contained misrepresentations about the pertinent reliability study - specifically, that the study cited in the Application was conducted on the first set of data, not the second set, which was the data ultimately used. The Application stated that the methodology used

[24] Although the hearing transcript indicates that Relator's Table was in the courtroom and was provided to the court and the Defendants during the hearing, the district court did not acknowledge the Table in its decision.

to manually draw image maps had demonstrated an inter-rater reliability coefficient between 0.94 and 0.99. According to Albert, such reliability numbers are "very high" and demonstrate consistency in results between raters, here Killiany and Gomez-Isla. Relator claimed that when the second set of data was tested for reliability, the Pearson Correlation coefficient dropped from 0.96 to 0.54.

The district court disregarded Jones's testimony on this issue. Based on Jones's experience as lead statistician for Core B, the district court thought that Jones was "likely . . . qualified to provide expert testimony regarding a reliability study," but nonetheless rejected his testimony because (1) "it is not clear . . . that the Relator has put himself forth as an expert consistent with [Federal Rule of Civil Procedure] 26(a)(2)" and (2) the Relator failed to provide a proper foundation upon which to accept his conclusions. The district court found Jones similarly deficient as a lay witness, finding that he did "not provide sufficient competent evidence of his personal knowledge," because (1) much of the data he received from Killiany's research came by way of another statistician on the project, and (2) he provided no evidence conveying "when and how the reliability study was conducted, who randomly selected the twenty-five subjects for the study, and who actually conducted the study."

Jones did, in fact, list himself as a non-retained expert in his Rule 26(a)(2) expert disclosure, see Fed. R. Civ. P. 26(a)(2), and, to form his opinion, relied on personal knowledge obtained as leader of the statistical Core and information provided by the Defendants during discovery, see Fed. R. Evid. 702. The district court abused its discretion in excluding from consideration Jones's reliability study testimony. See Alt. Sys. Concepts, Inc., 374 F.3d at 32. Moreover, the district court did not account for the apparently undisputed fact that no reliability study was conducted on the re-measurements. For example, Albert was asked during her deposition whether "anybody . . . conducted any reliability studies on the remeasurements?" She replied, "No . . . . We didn't have reliability data for -- we didn't have another rater available. We would have had to redo the reliability studies, and to me the critical thing . . . was that [the measurements] were accurate." As Teitelbaum points out, however, the question was not only one of accuracy, but also one of reliability, specifically whether another reliability study was necessary after Killiany's re-measurements. Albert thought not, saying, "I thought that we were following the guidelines, that they were the same guidelines established in the reliability study, and that we were applying them as best we could."

Teitelbaum, on the other hand, opined that "it was inappropriate to claim that a blinded reliability study had been

used in the generation of preliminary data when the data reported was not generated pursuant to the reported reliable methodologies" and had not been subject to a reliability study. Similarly, Jones maintains that after he "analyzed the impact of the altered data on both the reliability study and the reported volumetric data results," it became clear that the changes that Killiany made were responsible for the statistical significance of the reported data and that the altered data resulted in a vastly lower Pearson Correlation coefficient. Any technique modifications that Killiany made after discovering anatomical anomalies meant that he was no longer using the precise method Gomez-Isla had employed when she previously outlined ECs under the original methodology. The reliability numbers published in the Application conveyed the reliability between Killiany and Gomez-Isla under the original, unmodified approach to outlining the EC. Relator and his experts suggest that once Killiany deviated from this methodology, another reliability study should have been conducted and its results included in the Application.

There are substantial disputes here on the veracity of the reliability study data included in the Application. The district court erred in concluding otherwise.

3. Misrepresentation of Compliance with the PHS Terms and Conditions

Jones alleged that the Defendants misrepresented their compliance with the Public Health Services ("PHS") terms and

conditions, which outline investigation and reporting requirements when scientific misconduct is reported. Regulation 42 C.F.R. § 50.103(d)(1) requires each institution to "inquir[e] immediately into an allegation or other evidence of possible misconduct." The institution must contact outside authorities and report any situation in which, based on the initial inquiry, the institution determines that an investigation is warranted. 42 C.F.R. § 50.103(d). Jones claims that the inquiry Moss conducted was patently inadequate to satisfy the requirements of § 50.103(d). Relator also argues that Defendants were required to create a written record of the inquiry conducted about Killiany's alleged misconduct and report the results to the NIH Office of Research Integrity.

Noting that Relator had not properly pled his PHS terms and conditions certification claim in his Second Amended Complaint, the district court stated that Relator made this claim for the first time in his motion for summary judgment. Further, the district court stated that even if the claim were considered on the merits, it could not withstand summary judgment. The district court acknowledged that the "Applicant Organization" certification signed by the MGH Director of Grants and Contracts promised that MGH would comply with PHS terms and conditions, including 42 C.F.R.

§ 50.103.[25] It found, however, that the certification only promised compliance once a grant was awarded. Accordingly, because the alleged misconduct occurred before submission, it was not governed by the forward-looking certification.

We focus on the district court's procedural critique of Relator's pleading. In so doing, we agree with the court that it was not until Jones filed his motion for summary judgment that he propounded the theory that the Defendants' failure to investigate and report any inquiry was itself a false claim, independent of his claims regarding the use of falsified data not subject to a reliability study. Under Fleming v. Lind-Waldock & Co., 922 F.2d 20 (1st Cir. 1990), that was too late. Id. at 24 ("[I]nitial failure to satisfy the [pleading] burden in no way obligates the district court to allow the parties an opportunity to offer matters outside the pleadings. Simply put, summary judgment is not a procedural second chance to flesh out inadequate pleadings."); cf. Redondo Waste Sys. v. López-Freytes, 659 F.3d 136, 141 (1st Cir. 2011) (affirming denial of motion to amend where complaint did not sufficiently "apprise defendants of the claims against them,"

[25] The district court suggested that the certification signed by MGH was too vague to support what it styled an "express certification." We do not address this aspect of the district court's decision other than to note that we have since rejected the strict, categorical approach that it employed. See Hutcheson, 647 F.3d at 385-86.

which, the court stated, was "[t]he whole point of notice pleading").

In an effort to avoid this conclusion, Jones points to three paragraphs in the Second Amended Complaint that he argues demonstrate that his PHS terms and conditions compliance claim was properly pled:

> 22. On the October 2001 grant application, defendants Albert and Massachusetts General Hospital certified their assurances that the information supplied was "true, complete and accurate," that they accepted the obligation to comply with Public Health Service terms and conditions, and that they acknowledged liability under federal law for false or fraudulent statements or claims.
>
> 23. During the course of applying for and receiving NIH grant funds for the 2002-2007 funding cycle, defendants, and each of them, made or caused to be made statements and certifications in grant application forms, progress reports, vouchers, requests for progress payments and other writings necessary for the payment of federal funds.
>
> 27. On the basis of the false and fraudulent statements, defendants were awarded NIH grant funds. These funds were paid out and disbursed to defendants over time pursuant to the Public Health Service terms and conditions.

These statements do not contain any references to the Defendants' alleged violations of the PHS terms and conditions, namely, the failure to adequately investigate Jones's allegations of misconduct or report the Moss inquiry that was conducted in 2001. Moreover, read in context, paragraph 27's reference to

"false and fraudulent statements" does not refer to a certification of compliance with the PHS terms and conditions. Rather, it refers to statements set forth in the complaint (see ¶¶ 24-26) about the validity of the data, the blinding protocols used, and the inter-rater reliability coefficient applicable to the reported data.

The Second Amended Complaint also generally alleged that the Defendants "knowingly failed to take corrective action or disavow the false and fraudulent data after learning that their representations to NIH were false." But that allegation, too, read in context, does not aver an independent PHS terms and conditions certification claim. Instead, it relates to Albert's knowledge regarding the validity of Killiany's data and the claims made in the application based on that data. Thus, as the district court did, we conclude that any distinct, independent FCA claim resting on a statement of adequate investigation or an omission of proper reporting was not originally pled in the Second Amended Complaint. We affirm the district court's grant of summary judgment on this claim.[26]

[26] Although we affirm the district court's summary judgment determination on Jones's misrepresentation of compliance with the PHS terms and conditions claim, we do not suggest that the Moss inquiry is irrelevant as a factual matter to other aspects of this case. The nature and adequacy of the inquiry may still be relevant to Jones's other falsity claims.

C. Materiality

A false statement is material if it has "a natural tendency to influence, or [is] capable of influencing, the decision of the decisionmaking body to which it was addressed." Loughren, 613 F.3d at 307 (alteration in original) (quoting Neder v. United States, 527 U.S. 1, 16 (1999)) (internal quotation marks omitted). Jones claimed that the allegedly manipulated data relied upon and the reliability coefficient reported were material misrepresentations that would have a natural tendency to influence the Application reviewers.[27]

1. Manipulated Data

Because the district court failed to address Teitelbaum's and Jones's statements and other relevant record evidence, it did not make a materiality determination with regard to Relator's data manipulation claims. On the record before us, we conclude that it is likely that Relator's Table, Jones's testimony, and Teitelbaum's opinions about the manipulation of data - if credited - would be deemed material by a fact-finder. The allegedly false data produced the preliminary research results relied upon in the Project 3 proposal in the Application. If established, the notion that Killiany had been unblinded and had selectively manipulated

[27] Jones also claimed that the Defendants' allegedly false certification of compliance with the PHS terms and conditions was material. Because we affirm the district court's summary judgment determination on that issue, we do not address the materiality of that aspect of the certification.

data to produce a statistically significant result would certainly have had "a natural tendency to influence" the reviewers' evaluations. See id.

2. Reliability Correlation Coefficient

Having excluded portions of the reports from experts Schuff and Dávila-García, who alleged that the statements regarding the reliability study were material to NIH's decision to fund the Grant, the district court found that Jones "fail[ed] to satisfy the materiality element with respect to the statements concerning the reliability study."

a. Schuff's Materiality Determinations

The district court found Schuff unqualified to "testify as to the materiality of a statement regarding the NIH review process" because he did not "list any qualifications regarding the NIH application review process or the peer editing process." To the contrary, Schuff's curriculum vitae - submitted with his report - listed four NIA/NIH grant proposals on which he was a reviewer between 2006 and 2009 and numerous other experiences as a grant and peer reviewer for other institutions as well. Schuff's curriculum vitae also stated that he has specialized knowledge and training in relevant topics such as neuroimaging and neurodegeneration, has published more than 150 peer-reviewed articles in those and related fields, and has acted as the Principal Investigator on several clinical trials that utilized MRI to examine ROIs including the EC.

In light of the information in Schuff's curriculum vitae, we conclude that the district court abused its discretion by excluding Schuff's opinions regarding the materiality of the application statements discussing reliability. Thus, we treat those statements as if they had been admitted and consider them in our de novo review of the district court's summary judgment determination.

b. Dávila-García's Materiality Determinations

The district court similarly excluded Dávila-García's statement "that the reliability study was material to NIH's decision to fund the Grant." Although the court stated that Dávila-García "appears qualified to opine" on the materiality of statements in the Application concerning the reliability study, it rejected her report because it found that she "[did] not support her opinion with any evidence from the record." Specifically, the district court found that although Dávila-García stated that the reliability analysis was material because it was a required element of the application, it excluded her opinion because she "[did] not . . . provide any support for [her] statement from a statute, regulation, instruction manual, or . . . personal experience[,] . . . [and did not] cite any of the reviewers' comments from the Pink Sheets regarding the strengths and weaknesses of the Application." Further, the district court found that the record contradicted Dávila-García's testimony, because the Pink Sheets stated that "[t]he use of the Pearson correlation coefficients and

Student t-tests to assess reliability, as proposed, is inadequate." Despite the reviewers' disapproval of the proposed reliability methodology, they favorably evaluated the Application. Therefore, the Defendants argued, the reliability study could not have been material to the NIH's determination. The district court agreed.

The district court abused its discretion by concluding that Dávila-García did not sufficiently rely on personal experience in formulating her opinion about the materiality of the reliability study. If a witness relies primarily on experience, she must "explain how that experience leads to the conclusion reached, why that experience is a sufficient basis for the opinion, and how that experience is reliably applied to the facts." Fed. R. Evid. 702 advisory committee's note. Dávila-García had personal experience as a peer reviewer and was familiar with the NIH grant application process.[28] In her report, Dávila-García expressly stated that "each opinion [in the report] is based upon a reasonable degree of certainty, in light of [her] experience and background." She explained the peer-review process that grant applications undergo at the NIH:

> In making its ranking decision, NIH peer reviewers carefully consider several factors, as reflected in the NIH grant

[28] According to her curriculum vitae, Dávila-García reviewed four articles between 1999 and 2009, and acted as a grant reviewer on six committees between 2001 and 2010, including one NIH committee. Dávila-García also acted as a grant consultant on two projects, one with the NIH from 2007-2012.

> application. Instructions require applicants to include in the proposal a detailed research plan, describing the specific aims of the scientific project, preliminary data and progress reports on the work performed prior to the grant's submission, the procedures used and proposed to be used during the course of the project, and anticipated problems for the project with proposed solutions. These disclosures are essential to the peer-review process, as it is the responsibility of the reviewers to provide their best assessment of the chances for success and significance of that success on any given project.

She then applied her experience and knowledge of the peer review process to the facts in this case as she understood them:

> Within the grant proposal at issue in this case, the applicants make a number of statements that were material to the government's funding decision. Such material information undisputedly included the preliminary data and progress report of the scientists' ongoing research project, as well as the reliability of the blinded methodologies the scientists claimed they followed in validating that data . . . . Defendants represented that reliability studies had confirmed a high correlation between and among blinded raters (measuring in ranges between 0.94 and 0.99 for Pearson r correlation coefficients). Each of these statements was required to be made in the grant application, and each is fundamental to the peer review ranking of the application.

Moreover, we note that the Pink Sheet statement cited by the Defendants and the district court appeared in Section D of the Application, entitled "Research and Design Methods." Section D outlined the methodologies that would be employed in future studies to be conducted if grant funds were awarded. In contrast,

reliability study results from past studies were presented in Section C, "Progress Report/Preliminary Studies." Although we agree with the district court that the methodology concerns described in the Pink Sheets indicate that the Defendants' proposed method of future reliability testing was immaterial to the reviewers' decision, that conclusion does not satisfy the materiality determination about the results of the reliability study conducted in previous research projects. The alleged falsity in this case does not rest upon the adequacy of the reliability method to be employed in the future, but rather on results allegedly already obtained in a past reliability study and relied upon by the applicants in their proposal.

In sum, the evidence brought forth by Relator on the reliability issue generates an issue of fact regarding materiality, specifically, whether providing a reliability coefficient of 0.54, or stating that no reliability study was conducted on the measurements that gave rise to the scientifically significant results, would be capable of influencing the reviewers' decision.

D. Knowledge

The text of the FCA and our case law make clear that liability cannot arise under the FCA unless a defendant acted knowingly. See 31 U.S.C. § 3729(a); Hutcheson, 647 F.3d at 388. In the district court, Jones pointed to Relator's Table and Teitelbaum's expert report, among other record evidence, as proof

that the defendants knowingly created falsified data and used that data to support statements in the Application espousing the promise of research demonstrating the significant predictive value of the volume of the EC in determining who will later develop AD.[29] As noted, the district court abused its discretion by failing to consider Jones's statements and Teitelbaum's report at summary judgment.

Jones also alleged that the parties knowingly submitted a reliability coefficient from Killiany's study that did not incorporate his second set of data, the set from which the study's conclusions were drawn. The district court found that Jones failed to establish that the parties knew that they were submitting a "statement regarding the Pearson Coefficient [that] was inaccurate." The district court found that "the record [wa]s silent as to whether the reliability study was conducted on the second set of data and whether the reference to the Pearson Coefficient related to the first or second set of data."

As noted, the district court's conclusion on this matter ignores Defendant Albert's statement that she did not have an additional reliability study performed on the re-measurements because she was focused on the accuracy and did not have another rater available for re-measurements. See supra Part II.B.2. Furthermore, Killiany testified that he reviewed Gomez-Isla's

[29] The relevant statements are quoted supra Part II.B.1.a.

tracings sometime in 1996 or 1997 after the reliability study had been conducted. Killiany recognized that in tracing the scans upon which the reliability study was based, Gomez-Isla was "trying to apply the same working definition [that he] was trying to apply at the time." At the same time, Killiany testified that he continued tracing scans into 1998 or 1999, making revisions when necessary, and acknowledged that it is possible that he re-measured scans that were part of the initial reliability study well after that study had concluded.

In light of the foregoing arguments, and construing all facts in favor of Relator, we conclude that Jones generated a genuine issue of material fact as to whether the Defendants acted knowingly when allegedly making false representations in the Application.

III.

The dispute at the heart of this case is not about resolving which scientific protocol produces results that fall within an acceptable range of "accuracy." Nor is it about whether Killiany's re-measurements, the basis for the preliminary scientific conclusions reported in the Application, are "accurate" insofar as they fall within a range of results accepted by qualified experts. Rather, the essential dispute is about whether Killiany falsified scientific data by intentionally exaggerating the re-measurements of the EC to cause proof of a particular

scientific hypothesis to emerge from the data, and whether statements made in the Application about having used blinded, reliable methods to produce those results were true. If the jury should find that statements in the Application are false, they must also determine whether those statements were material and whether the Defendants acted knowingly in violating the FCA.

Because we conclude that genuine issues of material fact remain on these central issues, we vacate the district court's order and remand for further proceedings consistent with this opinion. Costs are awarded to the appellant.

So ordered.