Dear Richard:

You inquire about my analysis of your data and of the Hansen data. Neither was ever provided to you. Using well established methods I made multiple **fabrication** tests of your data. There was no evidence of **fabrication**. Drs. Carriquiry and Kaiser used complex methods for detecting **fabrication** recommended by the Government agency responsible for developing such methods and for overseeing their use in PHS agencies. They found no evidence of **fabrication**. I found the Hansen data were **plagiarized**, as later confirmed in Court. I found the Hansen data to be **falsified**, as later confirmed in Court. The law establishes three forms of data fraud: **fabrication, falsification,** and **plagiarism**. You were charged with **fabrication** and all the tests show there was no **fabrication**.

It may be best to provide some commentary on my statistical background. My prewar experience had been high school dropout to take a manufacturing production line job. It was the depths of the Depression. We were on welfare. Night school (Electrical Engineering, Georgia Tech) led to employment in the Electrical Engineering departments of a power company and then a telephone company. My professional involvement with statistics began with my first job upon returning from three years WWII Naval service. It was at Georgia Tech doing statistical analyses for corporate studies in industrial psychology in the Psychology Department, the beginnings of my involvement in psychology. The following year brought an appointment to the Mathematics faculty. In 1949-50 I became a student in a one-time applied statistics program at Yale, taught by the world's top statisticians as visiting professors. It was my good fortune to be assigned as a graduate assistant to Sir Ronald Fisher, universally regarded as the greatest statistician of all time. Not only was Fisher the Father of modern statistics, he was also the Father of modern population (quantitative) genetics which is how I got into neuro-behavioral genetics. Also on the visiting faculty were Frederick Mosteller and Philip Rulon of Harvard. Many regard Mosteller as the greatest statistician of the second half of the 20th century. Rulon held the Measurement chair at Harvard. In 1951 I went to Harvard as a post-doc with Mosteller and also worked in a Harvard affiliated research institute led by Rulon and American Association for the Advancement of Science President Kirtley Mather. There I was Project Director on two contracts, one in air traffic control for the Air Force, the other for simulator combat training for flag rank Naval officers. Next was a research consulting slot with the State of Connecticut for educational and labor market studies. I held various professional offices, most interesting being the Presidency of the Connecticut Chapter of the American Statistical Association. Connecticut had a high population of insurance statisticians (actuaries) as the Insurance State, of industrial statisticians (quality control engineers) as the high tech manufacturing center where mass production originated (clocks and arms), and of financial statisticians (accountants) as the leading commuter residential State for the New York banking industry. Two of my Executive Committee went on to Nobel Laureates in Economics (Tobin and Koopmans). I also served on an Institute of Mathematical Statistics Committee on Standards for Training of Statisticians. My career moved to academe in 1957 where I formally retired in 1986. I was named Distinguished Scholar at the University of Northern Iowa. I have been a regular reviewer for a number of scientific journals here and in Europe and for the National Institutes of Health and the National Science Foundation. After over two decades of retirement I have been accepting review requests less frequently.

You contacted me for advice on an indictment charging that “some of the data were fabricated” in a soy chip diet study of 60 research participants. More specifically you indicated it was known that some of the data were genuine but alleged later data were fabricated. I replied fabrication of data is a matter of great current interest in the financial community, the intelligence community, and the health research community. My advice was that you should contact the Office of Research Integrity to ascertain what, if any, assistance you could obtain from them. They were established as the Federal Agency

responsible for developing methods for detecting lack of integrity in research data and were touted in the statistical world for their contributions. They inherited some of the FBI experts in data fraud but early reports on formation of the ORI were not clear on the scope of their mission which was asserted to be Government wide on data fraud research and education but limited to PHS activities in investigatory authority. Your attorney did not see me as a potential witness and my suggestion was Dr Alicia Carriquiry who has a high reputation and who teaches forensic statistics at Iowa State University whose Statistical Laboratory has long been regarded as one of the top half dozen statistical institutes in the world.

You told me your attorney regarded statistics as worthless in the courtroom and any good lawyer could destroy statistical evidence. I commented my accountant brother who is operating vice-president of a financial house and on multiple boards of directors would be horrified to learn that any good lawyer could destroy the results of any audit. You indicated you were advised by your attorney that the judge held similar negative views of statistics. I am skeptical. My experience has been of lawyers trying to make statistics sound worthless only to have the judge chastise them with a lecture on statistics. My experience is not extensive but I have testified a few times. According to the Des Moines Register many years ago I was the witness who brought regression analysis into the judicial system as a standard method for assessing race and sex discrimination in wages and salaries. Some of the lawyers betrayed little competency in statistics. The judges I have encountered were more knowledgeable. When I expressed surprise once after trial at how much the judge knew he commented it was the job of judges to learn what they needed to know and he had obtained a crash education in statistics because he was the judge who heard the great redistricting case.

The difficulty with statistics is that a type of reasoning is required to which people are not accustomed. The fundamental basis of statistics is that the universe is governed by the laws of chance. The less scientifically educated can be misled, as your attorney suggests, by the fact the statistician will not say with certainty that something is or is not so. The statistician's work is based on the fact there is no certainty. That reality is expressed in the judicial system by the abstraction of levels of chance: “beyond a reasonable doubt” and “preponderance of the evidence”. There are studies on the levels of chance people ascribe to these terms. I have seen appeals court decisions remanding for failure to include the quantitative levels of probability in the court record. In the abstract we may identify a connection and prove *if A then B* but in the real world the exact proof is that *if A then B plus or minus e*. In popular parlance there is a margin of error. Statisticians are by the nature of their profession aware of error where most people are not. For example, people tend to think of computers as giving unquestionable calculations. However *A times B equals C* is actually *A times B equals C plus or minus e*. The margin of error is small but real. Forty years ago the National Bureau of Standards developed very complex algorithms for very simple arithmetic operations such as multiplication for the purpose of reducing that margin of error (*NBS Special Publication 339*, 1970). Other algorithms verified error levels in very complex calculations. I still use them occasionally and decry their absence from contemporary software packages.

The detection of research fraud rests on three basic scientific realities. The universe is governed by the laws of chance, hence we can test whether data follow the laws of chance or are fabricated. The phenomena of the real world result from many factors interacting with each other. The National Transportation Safety Board needs months to run down the specific factor or factors leading to a crash. The Mayo Clinic may run a hundred tests to discover why a body is not functioning properly and additionally consider their relationships to each other. Physiology and behavior vary statistically with differing genes and environment. To avoid detection the fraud perpetrator must be able to anticipate which tests and which interrelationships will be tested and design data which will pass those tests. The

third and never mentioned fact is that Pavlovian conditioning and operant conditioning were displaced by the discovery about half a century past that the human nervous system cannot manage ten concurrent concepts. Our air safety research revealed that airplane accidents stemmed from too much information—one can tell time more readily with a four number otherwise blank dial than with a face showing 60 tick marks. Weather maps went from detailed measures and locations to five or at most six-color displays. The keep-it-simple principle was born.
You sent me your data as being effects of a dietary soy supplement in a sample representative of U.S. adult males and females selected for obesity. It was alleged earlier participants were real but later ones were fabricated. This fitted the paradigm of standard industrial quality control. Quality control engineers test and statistically track products monitoring whether products show trends away from statistical expectations and specifications. Trends or deviations signal underlying production factors have changed leading the engineers to investigate to determine what changed and to correct the problem. The allegation that the underlying factors changed from dieter response to fabrication seemed a perfect fit. For three quarters of a century it has been conventional to display the statistics in the form of charts showing the sequential measurements and boundaries of expected margins of error. The methods originated with W. Edwards Deming (one of the Fathers of survey and census methods and the progenitor of Japanese manufacturing production and quality control methods) and with Walter Shewhart for whom the charting method is named. I tested your data and found no evidence of changes in the data, hence, no evidence of fabrication. For reasons cited above that fabrication is very difficult and because Shewhart charting has long been well established and successful as the basis for quality control I concluded there was no evidence of fabricated data. You forwarded my assessment to your attorney, Mr. Hansen.

Mr. Hansen wrote me it is impossible to tell whether data are fabricated on the basis of examining the data. (I was tempted to point out the recent major fraud cases in which the primary evidence was the CPA audits.) He indicated he could easily fabricate data so that it could not be detected. He indicated he would do so and send me a data set comparable to yours and challenged me to use my Shewhart methods to show his data were fabricated. He particularly emphasized that he had written an undergraduate thesis on Deming and fully understood the concept. As I recall there was an e-mail explicitly stating the issue was tthat earlier data were valid and the balance of the data were not.

I tested the Hansen data set as I would as a journal reviewer. I reported that the first three tests each showed data were falsified or, more precisely as a journal reviewer, they were not what they were represented to be. Specifically the results showed the data were not representative of the population to which inferences were to be made. For journal reviewing I would have stopped at that point, rejecting the manuscript and leaving it to the Editor to decide whether to investigate it as falsification or conclude the sampling procedures were defective.

As requested I did apply the Shewhart methods and reported to Mr. Hansen they showed no fabrication. Since he had clearly stated he understood the method would test whether some of the data were genuine and the balance fabricated, since he had prepared the data, and since the data tested as not being fabricated, it was evident he knew the data were not fabricated. It seemed impossible Mr. Hansen could have obtained such data elsewhere so the data must be falsified Fleming data. A plagiarism test was statistically significant in the range of seven orders of magnitude. In layman's terms the chances the Hansen data were not plagiarized from the Fleming data are less than one in ten million or of the same order of magnitude as the chances that one of the jurors will die in an automobile accident in the next 24 hours. I saw no need for further plagiarism tests. At the time I concluded Mr. Hansen's intent was to test my analysis of the Fleming data to see if I arrived at a different conclusion when I was led to believe the data were fabricated. We communicated no further.

In all I made nine fabrication tests on the Fleming data and nine on the plagiarized Hansen data. None of these 18 tests showed any evidence of fabrication. All three falsification tests showed the Hansen data had been falsified. The plagiarism test speaks for itself. Fabrication, falsification, and plagiarism are the three forms of health data fraud defined by statute.

The Carraquiry-Kaiser report represents a totally different approach than mine. It follows along the lines suggested by the Office of Research Integrity. The ORI is the Governments agency for developing best methods for detecting research misconduct which would seem to establish its methods as a Government established standard.

I note, inter alia, that the Carriquiry-Kaiser report speaks of difficulties with the Hansen report which also arose at Trial. My reading of the report is that the authors were puzzled by the Hansen report because they could find no evidence of fabrication when they were told the data were fabricated. They explicitly excluded falsification tests, justified by that information but which I regarded as something of a deficiency.

I do not have readily available all of the Shewhart charts nor the plotting programs which generated them. Attached are work copies not fully labeled. On the density plots the header is not a label but part of the computer program which generated the graphic.

In summary I audited the Fleming data using a number of standard industrial quality control tests to determine whether some of the data were genuine and some fabricated. There was no evidence of fabrication. I similarly audited the Hansen data finding no evidence of fabrication. I applied several tests to see if the data were representative of the defined population group. The Fleming data were. The Hansen data were not, suggesting falsification. A comparison test showed the Hansen data were plagiarized from the Fleming data. Falsification tests rest on the effects of a large number of underlying factors. Falsifying the numbers for a few of those factors alters little of the underlying factor effects. The assessment of no evidence for fabrication of research participants in the Hansen data simply provides a confirmation of lack of evidence of fabrication of research participants. The Carriquiry-Kaiser report/testimony represent an entirely different and more complex set of tests for fabrication following the recommendations for testing for fabrication of the Federal agency charged with developing and promulgating such testing methods. With their entirely different approach from mine they also found no evidence of fabrication of the Fleming data and confirmed that result with the Hansen data. For report and testimony they were asked to respond only to the charge of fabrication. They were not asked to for either falsification or plagiarism and did not do so.

As I said at the beginning: Using well established methods I made multiple **fabrication** tests of your data. There was no evidence of **fabrication**. Drs. Carriquiry and Kaiser used complex methods for detecting **fabrication** recommended by the Government agency responsible for developing such methods and for overseeing their use in PHS agencies. They found no evidence of **fabrication**. I found the Hansen data were **plagiarized**, as later confirmed in Court. I found the Hansen data to be **falsified**, as later confirmed in Court. The law establishes three forms of data fraud: **fabrication, falsification,** and **plagiarism**. You were charged with **fabrication** and all the tests show there was no **fabrication**.