Understanding the Statistical Analysis of Carbon Dating of the Shroud of Turin

by Robert A. Rucker, MS (nuclear)
Rev. 1, July 14, 2019
Reviewed by Kevin Schwinkendorf, PhD (nuclear)

Abstract

The statistical analysis of the carbon dating of the Shroud of Turin is not generally understood because of the technical nature of statistical analysis, both the mathematics and the terminology. Several individuals have done statistical analysis of the 1988 measured values since the 1989 report by Damon, et al. In all these analyses, it was mathematically determined that the samples cut from the Shroud in 1988 were heterogeneous. This means that the samples were not representative of the Shroud, so that the conclusion of the carbon dating, i.e. that the Shroud dates to 1260-1390, should be rejected, i.e. given no credibility. This is probably because the C^{14} to C^{12} ratios of the samples were altered by neutron absorption which produced new C^{14} atoms on the Shroud. Thus, the conclusion of the 1988 carbon dating of the Shroud that the Shroud dates to 1260-1390 resulted from an inadequate statistical analysis that failed to recognize that the samples were heterogeneous. The effect of random errors, systematic bias, and sampling on scientific measurements are also discussed and applied to why the 1260-1390 carbon date should be rejected.

1. Summary

In 1988, samples were cut from the Shroud of Turin and carbon dated to a range of 1260 to 1390 AD, based on the statistical analysis of the measurements reported in Damon, et al (Ref. 1). This date of 1260-1390 conflicts with other evidence that the Shroud probably is from the time of Jesus (Section 6 of Ref. 10). This conflict has motivated several authors to perform additional statistical analysis on the 1988 data (Ref. 2 to 12). A recent paper by T. Casabianca, et al, titled "Radiocarbon Dating of the Turin Shroud: New Evidence from Raw Data" (Ref. 13) with supplementary material (Ref. 14) documents an important statistical analysis of the 1988 carbon dating of the Shroud which includes new data from the British Museum obtained by several Freedom of Information Act (FOIA) requests. Unfortunately, people not familiar with the mathematics or terminology used in statistical analysis will probably not properly understand important concepts in these documents. Perhaps of greatest importance, the authors of these statistical analyses (Ref. 2 to 14) agree that the samples sent to the three laboratories in 1988 were heterogeneous, which is also referred to as non-homogeneous, which means that:

- The samples cut from the Shroud were not representative of the average composition of the Shroud.
- This is because the samples were essentially different from each other relative to how they were measured in carbon dating, even though they were next to each other before being cut from the Shroud.

- Since carbon dating is done by measuring the C^{14} to C^{12} ratios in the samples, this means the three samples from the Shroud contained statistically different C^{14} to C^{12} ratios.
- Which means something had changed the C^{14} to C^{12} ratios in the samples, with the amount of this change being unknown for any given sample.
- Which means each measurement, though presumably measuring the correct C^{14} to C^{12} ratio in each sample, could have obtained dates that were different than the true date by an unknown amount.
- Which means that the uncorrected average value of 1260 ± 30 could be wrong by an unknown amount.
- Which means the uncorrected date of 1260 ± 30 should be rejected as not valid.
- Which means the corrected date range of 1260 to 1390 AD has no basis, and thus should also be rejected, i.e. given no credibility, because it could be wrong by an unknown amount.

Thus, the dates obtained during the 1988 carbon dating of the Shroud ought to be rejected, i.e. given no credibility, because the samples were heterogeneous, also called non-homogeneous, because they were not representative of the entire Shroud. The date range of 1260 to 1390 AD was accepted for the Shroud due to an inadequate statistical analysis in Damon (Ref. 1) that failed to recognize that the samples were heterogeneous, i.e. not representative of the entire shroud. The samples were heterogeneous evidently because a systematic error or bias had altered the C^{14} to C^{12} ratios for the samples as a function of the distance from the bottom of the cloth.

2. Errors in scientific measurements

All measurements, including carbon dating, are affected by things that cause the measured values to be in error to various degrees. There are two types of measurement error: random errors and systematic errors. All measurements are affected by random processes that can cause the measurements to be a little high one time or a little low another time. The effect of these randomly positive or negative errors can be minimized by doing many measurements so when the average value is calculated, the positive errors and the negative errors would cancel each other to a large extent. While all measurements are affected by these random errors, measurements, including carbon dating, can sometimes also be affected by things that cause systematic errors, also called a systematic bias. This type of error is called “systematic” in contrast to “random” because it can be caused by something that results in the measurements being in error in only one direction, either positive or negative. This means the effect of a systematic error does not cancel out when many measured values are averaged.

A simple example of a systematic error would be measuring the distance between two points with a ruler that is 5% shorter than the standard 12 inches. This shorter ruler would cause the measured value to be 5% higher than the true value, for example a measured value of 105 feet between two points instead of the true value of 100 feet. This 5% difference between the measured value and the true value is the systematic error or bias. No matter how many times the measurement is performed, this 5% error would be present in the average value. A more detailed example of a series of measurements being affected by a random error and a systematic error or bias, and the resulting statistical analysis, is given in Section 7 of Ref. 10. This example should be carefully studied to fully understand the effect of a systematic error, and how the presence of a systematic error can be detected by a statistical analysis of the measurement data.

Not only does a systematic error not cancel out with many measurements, it can also cause a significant error in the measured values. For example, since carbon dating of cloth is performed by measuring the amount of C^{14} in a sample of the cloth relative to C^{12}, a cloth from the time of Jesus, about 33 AD, would be carbon dated to 1260 AD if the C^{14} to C^{12} ratio in the sample were increased by only 16%. This large change in the date is caused by a small change in the C^{14} to C^{12} ratio because the half-life of C^{14} is 5730 years. If a systematic error is affecting the measurements, its existence needs to be recognized and the magnitude of the error needs to be quantified if possible so the measurements can be corrected. If there is a systematic error affecting the measurements, and if this systematic error is not quantified to permit correction of the measured values, then the measured values could be in error by an unknown amount. If this is the case, then the measured values should be rejected, i.e. have no credibility. This applies to the carbon dating of the Shroud.

To determine if a systematic error or bias is affecting a series of measurements, a statistical analysis is performed on the measured values in comparison to the measurement uncertainties. If the range of the measured values has a very low probability of occurring, given the measurement uncertainties, then a systematic bias is probably affecting the measurements. For example, if samples of a material are sent to three different laboratories, and each laboratory measures its sample multiple times, each laboratory can calculate an average value for its sample. If the range or difference in these average values is much larger than should occur due to the random errors resulting from the measurement uncertainties, then the presence of a systematic error is the only other option to explain the excessive range in the average values from the three laboratories. The presence of such a systematic error is the result of something causing the measurements for the samples to be fundamentally different. It should not be automatically assumed that the measurement personnel, procedures, materials, or equipment are at fault. If there is no convincing evidence that these are at fault, then the samples are probably fundamentally different for some reason. Something has changed the characteristic that is being measured in the samples.

3. The Problem of Sampling

A related problem is proper sampling. Consider a fixed large steel container holding multiple compounds mixed in liquid. If you want to determine the contents in the container, one way would be to analyze all the contents in the container. But this is usually not practical. The analysis process is often destructive so that the entire contents of the container would be altered or destroyed, thus leaving none of the material for later use, or the container may be too large to analyze all the material. For this reason, a small sample would usually be removed from the container and analyzed to determine what is in the container. But for this to yield the correct results, the sample removed from the container must be representative of all the material in the container, i.e. the sample must have the average composition of the material in the container. If this is true for a sample, then the sample is called homogeneous, where “homo” means being the same. If the sample is not the same as the average composition in the container, then the sample is called heterogeneous, where "hetero" means being different. A heterogeneous sample is often also called non-homogeneous. A correct scientific analysis of a heterogeneous sample of the material in the container would not yield correct information about the contents of the container because the sample did not have the average composition of material in the container.

For example, if the liquid was water and some of the compounds in the container had very low solubility in water, then they would mostly settle to the bottom of the container as a precipitate. Taking a sample from the center of the container would not sample from this precipitate at the bottom of the container, so the sample would not properly represent all the material in the container. It would be a heterogeneous sample. However, the person doing the sampling of the container could easily assume that it was a homogeneous sample. He would only recognize it as a heterogeneous sample if he were to take multiple samples from various elevations within the container, with the results changing as a function of (depending on) the vertical location from which the sample was taken. Since the analysis of each sample would include the effects of random errors in the measurements, he would only recognize that the results were a function of the vertical location in the container if a statistical analysis of the measurements indicated that the range of the measured results were larger than expected given the measurement uncertainties produced by the effects of the random variations in the measurements. If this were the case, the variation in the measurements as a function of the vertical height in the container would be attributed to a systematic bias resulting from the settling of material in the container.

Thus, in practice, samples are recognized as homogeneous when a statistical analysis of the data indicates that the variation in the measured values can be explained by the measurement uncertainties, i.e. that to a high probability the measured values are consistent with the measurement uncertainties. If the statistical analysis indicates that the probability is very low that the range or distribution of the measured values is consistent with the measurement uncertainties, then the only other option ought to be recognized as probable, i.e. that a systematic error or bias has altered the measured values. In this case the samples are not "homogeneous", so they are called "non-homogeneous" or "heterogeneous". In the concepts of statistical analysis, homogeneous samples are said to come from the same population of values whereas heterogeneous samples are said to come from different populations of values.

In simple terminology, samples are "homogeneous" if they are essentially the same (measured values consistent with the measurement uncertainties) in the quantity being measured, i.e. the C^{14} to C^{12} ratio for carbon dating. Multiple samples taken from the same piece of cloth or from nearby areas of the same rock strata ought to be "homogeneous", i.e. essentially the same. And in simple terminology, samples are "heterogeneous" or "non-homogeneous" when they are essentially different (measured values not consistent with the measurement uncertainties) in the quantity being measured. For samples that are homogeneous, random errors can account for the differences in the measured values, and no systematic error is implied. But for heterogeneous (non-homogeneous) samples, random errors alone cannot account for the differences in the measured values, so the presence of a systematic error is implied.

From the previous discussion, a systematic error can cause the measured values to be wrong to an unknown amount. Thus, measurement of heterogeneous (non-homogeneous) samples can be wrong to an unknown amount due to the presence of a systematic bias. Measured values from heterogeneous samples should be rejected from consideration because they could be wrong by an unknown amount. The purpose of doing a statistical analysis of the data is not only to determine the best average or mean value that should be reported but also to determine whether a systematic error could have affected the measurements, in which case the measured values should be rejected, i.e. given no credibility. If the statistical analysis reveals an inconsistency between the measured values and the measurement uncertainties, the uncertainties should never simply be ignored (as was done in Damon, et al), for this could hide the presence of a systematic bias that could invalidate the measured values.

It should not be implied by the above discussion that the carbon dating equipment used in the 1988 carbon dating of the Shroud was not measuring the correct C^{14} to C^{13} ratio in the samples. In fact, it is believed that the carbon dating equipment was measuring the correct C^{14} to C^{13} ratio that was in the samples. The problem was not with the measurements but with the samples. The C^{14} to C^{13} ratios in the samples were probably changed by neutron absorption in the Shroud which produced new C^{14} in the samples, which was correctly measured by the carbon dating equipment. It is believed that this neutron absorption increased the amount of C^{14} in the samples by 16% and the scientific equipment correctly measured this amount of C^{14} in the samples. This increase in the C^{14} in the samples shifted the calculated date from the time of Jesus (about 33 AD) to the carbon date of 1260.

4. 1988 Carbon Dating and 1989 Statistical Analysis

In 1988, samples were cut from one corner of the Shroud and sent to three laboratories at Tucson, Zurich, and Oxford for carbon dating. As reported in Damon, et al (Ref. 1), the averaged value for the four measurements at the Tucson laboratory was 646 ± 31, the average value for the five measurements at the Zurich laboratory was 676 ± 24, and the average value for the three measurements at the Oxford laboratory was 750 ± 30. These dates are given in years before 1950, and correspond to AD dates of 1304 ± 31, 1274 ± 24, and 1200 ± 30. (Tables 1 and 2 of Ref. 11). The reported average of these three values in Damon was 1260 ± 31 AD (1260 AD with a 1-sigma uncertainty of 31 years). This is the raw or uncorrected value. When this value was corrected for the changing amount of C^{14} in the atmosphere, a range of 1260 to 1390 was obtained. That the bottom limit of this range is the same as the uncorrected value, both being 1260, is only fortuitus, i.e. occurs by chance.

However, as discussed above, measured values cannot simply be assumed to be correct. Prior to being accepted as correct, the measured values must pass a statistical analysis of the measured values relative to the measurement uncertainties to assure variations in the measured values are only due to random measurement errors and not a systematic error. This is because the presence of a systematic error could cause the measured values to be wrong by an unknown amount. This is an important point that is not usually understood by a layman. In the statistical analysis reported in 1989 (Damon, et al, Ref. 1), they found that the range of the measured values was not consistent with the measurement uncertainties. Their solution to this inconsistency was to assume that the measurement uncertainties were in error, with no further justification given. This effectively assumed that a systematic error could not be affecting the measurements. This assumption in Damon, et al, was the root cause that led to the faulty interpretation of the measured values, i.e. that the Shroud dates to 1260-1390. This existence of this assumption was not readily apparent due to the technical terminology used in a journal article.

Since the statistical analysis in Damon did not prove a systematic bias had not affected the measurements, the average value of the three laboratory average values should not have been accepted as necessarily correct. In statistical analysis terminology, this raw or uncorrected result of 1260 ± 31 should have been "rejected" from use in dating the Shroud because it could have been wrong by an unknown amount due to the possible presence of a systematic error. But when the measurement uncertainties were assumed to be in error so they could be ignored, no statistical analysis was then possible, and the 1260 ± 31 value became the claimed result. This 1260 ± 31 value was then corrected for changes in the amount of C^{14} in the atmosphere over time thus resulting in the range of 1260 to 1390. The 1260 to 1390 range is stated in Damon to be a two-sigma range, which means there should be a 95% probability that the true date for the Shroud falls within this range. This 95% probability range of 1260 to 1390 was used by the managers of the laboratories and the media to prove that the Shroud was made in the 13th or 14th centuries and thus could not be the authentic burial cloth of Jesus. But this supposed 95% probability range of 1260 to 1390 was based on an average value of 1260 ± 31 that should have had no credibility. It should have been rejected because it failed to pass the statistical analysis test of consistency between the range of the measured values and the measurement uncertainties. The range of the measured values was not consistent with the measurement uncertainties because the measurements of the C^{14} to C^{12} ratios in the samples not only included random measurement errors but also very probably something that had systematically altered the samples causing a systematic error or bias in the reported results. There is only about a 1.4% probability that the range of the 1988 carbon dates was consistent with the measurement uncertainties (significance level =1.39% for material 1, lower left corner of Table 2, reproduced from Table 6 of Ref. 11).

The presence of a systematic error in the measurements of the C^{14} to C^{12} ratio could cause the measurements to be wrong by an unknown amount, and thus the calculated date to be wrong by an unknown amount. Because this probable systematic bias was not recognized or quantified, the date of 1260 AD could be wrong by any amount, and should thus have been rejected. Their <u>assumption</u> that the measurement uncertainties were in error allowed them to <u>assume</u> the 1260 ± 31 value was correct, which then produced their claimed 95% confidence range of 1260 to 1390. This is how the reported average value of 1260 ± 31 that should have had no credibility became perceived as a 95% confidence range of 1260 to 1390 supposedly with full scientific backing.

The carbon dating of the Shroud in 1988 and the statistical analysis of the data in 1989 (Damon) significantly decreased research for about a decade. But during this time, researchers continued to evaluate options to properly understand the Shroud. Based on a broad consideration of the evidence, most researchers gradually concluded that the Shroud was very probably from the first century and its carbon dating to 1260 to 1390 AD must therefore be in serious error (Ref. 15 and 16). Various hypotheses were considered to explain this discrepancy, i.e. the difference between about 33 AD and 1260 to 1390 AD. It is unlikely any contamination hypothesis is the correct explanation because the amount of contamination on the samples would have to be extremely high (60% to 80% of the weight of carbon in the sample) and because various types of severe cleaning of the Shroud samples was performed in 1988 without any apparent effect on the date (Ref. 20). The two leading hypotheses to explain this error between about 33 AD and 1260-1390

AD are the invisible reweave hypothesis and the neutron absorption hypothesis. According to the invisible reweave hypothesis, this error resulted from the interweaving of new material into the threads of the old linen, but significant objections have been raised against this hypothesis (section 2 of Ref. 12 and chapter 9 of Ref. 16). According to the neutron absorption hypothesis (Ref. 12), this error resulted from neutrons included in the burst of radiation emitted from the body that caused the image on the Shroud (Ref. 17 to 18). A small fraction of these neutrons would have been absorbed in the trace amount of N^{14} in the linen to produce new C^{14} in the cloth by the [N^{14} + neutron $\rightarrow$ C^{14} + proton] reaction. This would have increased the amount of C^{14} on the samples by 16%, which would have shifted the carbon date from about 33 to 1260 AD.

5. Understanding the Statistical Analysis in Casabianca, et al

Table 1 in this paper reproduces Table 1 in Casabianca, et al, (Ref. 13). This table summarizes the new data obtained from the British Museum regarding the 1988 carbon dating of the Shroud. With over 700 pages being obtained from the British Museum (Ref. 13), a significant amount of time will probably be needed to fully understand and organize this new information so further analysis and conclusions will probably be forthcoming. This new information is very important and should be included in future statistical analysis. The statistical analysis in Casabianca, et al, is so technical that an evaluation of it will not be attempted. But the conclusions in Casabianca will be considered to facilitate people's understanding. The statistical analysis in Casabianca arrives at conclusions that are generally consistent with previous statistical analysis by Remi Van Haelst (Ref. 2 to 5), Bryan J. Walsh (Ref. 6 and 7), A.C. Atkinson, Giulio Fanti, Fabio Crosilla (Ref. 8 and 9), and the author (Ref. 10 to 12). It is also consistent with the discussion in the previous sections of this paper. Statements in the section labelled "Discussion" on pages 6 to 8 of Casabianca are discussed below.

"The conclusions of the various statistical analysis methods applied to both the *Nature* and raw results *intra* and *inter* laboratories are concordant." This means a statistical analysis of the new (raw) data listed in Table 1 gives consistent (concordant) results to a statistical analysis of the data published in Damon whether the analysis is done by comparing the differences in the data between (inter) the three laboratories or by comparing the differences in the data obtained by (intra) each laboratory separately. Thus, statistical analysis of the new data from the British Museum produces the same basic conclusions as an analysis of the older data in Damon. However, this does not mean the conclusions in Damon are correct – see below.

Regarding Table 1, the raw data obtained from the British Museum confirm that eight measurements were performed at the laboratory in Tucson, Arizona, rather than just the four values reported in Damon. Why did those doing the statistical analysis think it necessary and justifiable to collapse the eight measurements performed into the four values reported? Perhaps it was that this process eliminated the highest and lowest of the Arizona values, thus diminishing the inconsistency between the range of the measurement data and the measurement uncertainties. The measured values did not change between the raw data and the data published in Damon except for the last two reported measurements by Zurich. The cause of this change is not understood. But many of the measurement uncertainties were changed from the raw data to the data published in Damon: Arizona (measurements 6 and 7, 676 ± 40 and 540 ± 37 combine to give 608 ± 27 instead of the value in Damon = 606 ± 41), for Oxford the uncertainties change as follows (53 → 65, 30 → 45, and 46 → 55), and for Zurich as follows (47 → 45, and 46 → 51). All but one of these changes is an increase in the measurement uncertainty. The result of these changes in the uncertainty is again to diminish the inconsistency between the range of the measurement data and the measurement uncertainties.

"The two modified radiocarbon dates which were achieved using the same standards, were clearly not identical within errors." This apparently refers to measurements 5 and 6 by Arizona where the raw 1 data lists 676 ± 40 and 540 ± 37. The difference between these two values is $136 \pm (40^2 + 37^2)$^0.5 = 136 ± 54, which indicates these two measurements are statistically different at the 136/54 = 2.5 sigma level.

"Our statistical results do not imply that the medieval hypothesis of the age of the tested sample should be ruled out." While it is true that the statistical analysis does not rule out the possibility of the medieval hypothesis being true, it is also true that the statistical results cannot prove the medieval hypothesis is true. This is because for heterogeneous samples, the C^{14} to C^{12} ratio in each sample could have been changed by an unknown amount. In this case, even if the C^{14} to C^{12} ratio in each sample is measured accurately, the resulting calculated dates would also have been changed by unknown amounts. Thus, for heterogeneous samples, no reliable conclusions can be reached other than to reject the data. The probability that the three laboratory's average values are consistent with each other, given the measurement uncertainties stated in Damon, which means that the samples are not heterogeneous because there is not a systematic bias affecting the measurements, which means that the medieval hypothesis is true, is calculated to be 1.4% by a Chi-squared statistical analysis (significance level =1.39% for material 1, lower left corner of Table 2, reproduced from Table 6 of Ref. 11).

"Each TS raw and published radiocarbon date indicates a medieval interval for the fabric. Nevertheless, this reasoning would simply assume a constant amount of ^{14}C atoms among the subsamples. This basic assumption is not supported by the heterogeneity of the TS raw data …". This says the "basic assumption" is that the samples contain the same amount of C^{14}, so that the samples are homogeneous (essentially the same). But this assumption is not supported by the TS (Turin Shroud) raw data. Rather, the data indicates the samples are heterogeneous (essentially different) though they were located next to each other on the Shroud. Thus, the results cannot be trusted to be accurate and should be rejected from use in dating the Shroud.

"The hypothesis of a statistical significance only due to some difference in measurements among the laboratories is weakened by the fact that the results were correct and consistent for the three control samples …" This argues the problem is not with the measurement equipment, materials, or procedures used in measuring the samples from the Shroud. This is because dates were measured reasonably accurately for three other fabrics at the same time as the Shroud samples. These three other fabrics are listed in Table 2: material 2 = line from the tomb of Qasr Ibrim in Egypt, material 3 = mummy of cleopatra from Thebes in Egypt, and material 4 = cope of St. Louis d'Anjou of France. This evidence indicates the C^{14} to C^{12} ratio was correctly measured in each Shroud sample, but the C^{14} to C^{12} ratio had somehow been increased in each sample. It is believed this increase in the C^{14} to C^{12} ratio probably resulted from neutron absorption producing additional C^{14} on the Shroud.

There was a "significant statistical trend of the TS raw data" that "showed a significant decrease in the radiocarbon age as one gets closer to the centre of the sheet (in length, from the tested corner). This variability of the *Nature* radiocarbon dates in a few centimeters, if linearly extrapolated to the opposite side of the TS, would lead to a dating in the future." This slope or gradient in the radiocarbon age (Figure 3 of Ref. 11) indicates the magnitude of the systematic bias is dependent on the original position of the sample on the Shroud, with older dates toward the bottom of the cloth and more recent dates as the location is moved away from the bottom of the cloth. In the invisible reweave hypothesis, this slope in the age is explained as due to the changing fraction of new vs. old material in the reweave as a function of location. In the neutron absorption hypothesis (Ref. 12), this slope in the age is explained by the natural distribution that neutrons would have taken in the tomb after being emitted from within the body as part of the radiation burst that formed the image. The reference to the possibility of dates into the future is consistent with the neutron absorption hypothesis (Figures 11 to 14 of Ref. 12), which predicts most locations on the Shroud should carbon date to the future (assuming the same equations were used) due to neutron absorption producing more C^{14} in the linen than would normally be present.

In the last paragraph, "the presence of serious incongruities among the raw measurements … strongly suggest that homogeneity is lacking in the data." Since the samples are not homogeneous, a systematic bias is evidently present in the measured values which could have changed the values by an unknown amount. This means the measured data should be rejected from any degree of credibility.

In the last sentence, "It is not possible to affirm that the 1988 radiocarbon dating offers 'conclusive evidence' that the calendar age range is accurate and representative of the whole cloth." Since this analysis does not affirm "conclusive evidence" for the 1260 to 1390 date for the Shroud, the conclusion in Damon (Ref. 1) that "the linen of the Shroud of Turin is mediaeval" ought to be rejected, consistent with everything said above.

6. Conclusion

In Ref. 2 to 14, the eleven authors who performed statistical analysis of the 1988 carbon dating of the Shroud concluded the samples sent to the three laboratories were heterogeneous, i.e. statistically different from each other regarding their C^{14} to C^{12} ratios. Since the samples sent to the laboratories were next to each other on the Shroud, something evidently had changed this ratio for the different samples. Thus, the average value from the three laboratories (1260 ± 31) should be rejected as not valid, which leaves no basis for the 1260 to 1390 AD range. The conclusion in Damon that "The results provide conclusive evidence that the linen of the Shroud of Turin is mediaeval" ought to be rejected, i.e. given no credibility. Casabianca, et al (Ref. 13 and 14) is not saying the 1988 carbon dating is essentially correct but should have a wider uncertainty range. Rather, Casabianca, et al, confirms the previous statistical analysis (Ref. 2 to 12) that the samples are heterogeneous, so the uncorrected date (1260 ± 31) and the corrected range (1260 to 1390 AD) ought to be rejected from use in dating the Shroud.

It is believed that the C^{14} to C^{12} ratio in each Shroud sample was accurately measured in 1988 because samples of three other fabrics (Table 2) that were dated at the same time as the Shroud samples were dated reasonably accurately. Thus, the problem was not with the measurements but with the samples. The C^{14} to C^{12} ratios for the Shroud samples had evidently been changed. The best explanation for this is probably the neutron absorption hypothesis (Ref. 12). This concept hypothesizes that neutrons were part of the burst of radiation emitted from the body that formed the image (Ref. 17 and 18). These neutrons caused new C^{14} to be produced in the samples primarily by neutron absorption in the trace amounts of N^{14} in the linen by the [N^{14} + neutron → C^{14} + proton] reaction. The C^{14} content must be increased by only 16% to shift the carbon date from 33 AD to 1260 AD. The four known characteristics related to carbon dating of the Shroud are the date, slope, and range of the data obtained in the 1988 carbon dating of the Shroud, and the carbon dating of the Sudarium to 700 AD, since the Sudarium is believed to be related to the Shroud. The neutron absorption hypothesis (Ref. 12) is the only hypothesis consistent with all four of these characteristics, but predictions of the neutron absorption hypothesis have not yet been tested (Ref. 19).

References

1. P. E. Damon, et al. (21 authors), "Radiocarbon Dating of the Shroud of Turin" in *Nature,* Vol. 337, No. 6208, pages 611-615, February 16, 1989.
2. Remi Van Haelst, "Radiocarbon Dating the Shroud, A Critical Statistical Analysis", 1997
3. Remi Van Haelst, "Radiocarbon Dating the Shroud of Turin, The Nature Report", June 1999
4. Remi Van Haelst, "A critical review of the radiocarbon dating of the Shroud of Turin. ANOVA – a useful method to evaluate sets of high precision AMS radiocarbon measurements", June 1999
5. Remi Van Haelst, "Radiocarbon Dating the Shroud of Turin, A critical review of the Nature report (authored by Damon et al) with a complete unbiased statistical analysis", October 2002
6. Bryan J. Walsh, "The 1988 Shroud of Turin Radiocarbon Tests Reconsidered, Part 1", 1999
7. Bryan J. Walsh, "The 1988 Shroud of Turin Radiocarbon Tests Reconsidered, Part 2", 1999
8. Marco Riani, A. C. Atkinson, Giulio Fanti, and Fabio Crosilla, "Carbon Dating of the Shroud of Turin: Partially Labelled Regressors and the Design of Experiments", May 4, 2010
9. Marco Riani, A. C. Atkinson, Giulio Fanti, and Fabio Crosilla, "Regression Analysis with Partially Labelled Regressors: Carbon Dating of the Shroud of Turin", Journal of Statistical Computation and Simulation, 23:551-561, 2013
10. Robert A. Rucker, "The Carbon Dating Problem for the Shroud of Turin, Part 1: Background", July 7, 2018.
11. Robert A. Rucker, "The Carbon Dating Problem for the Shroud of Turin, Part 2: Statistical Analysis", July 7, 2018
12. Robert A. Rucker, "The Carbon Dating Problem for the Shroud of Turin, Part 3: The Neutron Absorption Hypothesis", July 7, 2018
13. T. Casabianca, E. Marinelli, G. Pernagallo, and B. Torrisi, "Radiocarbon Dating of the Turin Shroud: New Evidence from Raw Data", accepted February 15, 2019 for publishing in *Archaeometry*.

14. "Supplementary Material" for the above paper by Casabianca, et al.
15. Mark Antonacci, "The Resurrection of the Shroud", 2000, 328 pages, M. Evans and Company, Inc., ISBN 0-87131-890-3
16. Mark Antonacci, "Test the Shroud", 2015, 502 pages, LE Press, LLC, ISBN 978-0-9964300-1-2
17. Robert A. Rucker, "Information Content on the Shroud of Turin", June 2016.
18. Robert A. Rucker, "Role of Radiation in Image Formation on the Shroud of Turin", October 5, 2016.
19. Robert A. Rucker, "Testing the Neutron Absorption Hypothesis", April 8, 2019
20: Robert A. Rucker, "Carbon Dating of the Shroud of Turin to 1260-1390 AD is Not Explained by Normal Contamination", July 10, 2019

Biography

Robert A. Rucker earned an MS degree in nuclear engineering from the University of Michigan and worked in the nuclear industry for 38 years primarily in nuclear reactor design, nuclear criticality safety, and statistical analysis for quality control of nuclear material inventories. He holds Professional Engineering (PE) certificates in nuclear engineering and in mechanical engineering. He organized the International Conference on the Shroud of Turin (ICST-2017) held July 19-22, 2017, in Pasco, Washington. His papers can be downloaded from the research page of his website at http://www.shroudresearch.net/research.html. Send comments, questions, or corrections to robertarucker@yahoo.com.

Table 1. Radiocarbon Dates Before Present (1950) of the Turin Shroud (± error)

Arizona Raw 1	Arizona Raw 2	Arizona Nature	Oxford Raw	Oxford Nature	Zurich Raw	Zurich Nature
606 ± 41	606 ± 41		795 ± 53	795 ± 65	733 ± 61	733 ± 61
574 ± 45	574 ± 45	591 ± 30	730 ± 30	730 ± 45	722 ± 56	722 ± 56
753 ± 51	753 ± 51		745 ± 46	745 ± 55	635 ± 57	635 ± 57
632 ± 49	632 ± 49	690 ± 35			617 ± 47	639 ± 45
676 ± 40	676 ± 59				595 ± 46	679 ± 51
540 ± 37	540 ± 57	606 ± 41				
701 ± 47	701 ± 47					
701 ± 47	701 ± 47	701 ± 33				

Table 2. Recalculated Statistical Analysis with 8 Tucson Values

	Material 1	Material 2	Material 3	Material 4
Source of material:	Shroud of Turin	Linen from tomb at Qasr Ibrim, Egypt	Mummy of Cleopatra from Thebes, Egypt	Cope of St. Louis d'Anjou of France
Expected date:		11th to 12th Century AD	110 BC to 75 AD	1290 to 1310 AD
Laboratory	**Individual Measurements of C^{14} Date, Years Before Present (YBP, Present = 1950)**			
Tucson, Arizona	606 ± 41	922 ± 48	1838 ± 47	724 ± 42
	574 ± 45	986 ± 56	2041 ± 43	778 ± 88
	753 ± 51	829 ± 50	1960 ± 55	764 ± 45
	632 ± 49	996 ± 38	1983 ± 37	602 ± 38
	676 ± 59	894 ± 37	2137 ± 46	825 ± 44
	540 ± 57			
	701 ± 47			
	701 ± 47			
Zurich, Switzerland	733 ± 61	890 ± 59	1984 ± 50	739 ± 63
	722 ± 56	1036 ± 63	1886 ± 48	676 ± 60
	635 ± 57	923 ± 47	1954 ± 50	760 ± 66
	639 ± 45	980 ± 50		646 ± 49
	679 ± 51	904 ± 46		660 ± 46
Oxford, England	795 ± 65	980 ± 55	1955 ± 70	785 ± 50
	730 ± 45	915 ± 55	1975 ± 55	710 ± 40
	745 ± 55	925 ± 45	1990 ± 50	790 ± 45
Laboratory	**Weighted Mean C^{14} Dates (YBP) Based on Above Values**			
Tucson, Arizona	646.52 ± 17.18	927.44 ± 19.70	1995.23 ± 19.89	721.67 ± 20.42
Zurich, Switzerland	676.14 ± 23.74	940.60 ± 23.16	1939.81 ± 28.47	685.16 ± 24.63
Oxford, England	749.17 ± 30.70	937.88 ± 29.43	1977.05 ± 32.71	755.76 ± 25.66
	Analysis of Interlaboratory Scatter			
Unweighted mean of unweighted means (YBP)	695.38 ± 32.15	937.33 ± 6.26	1968.82 ± 14.74	732.16 ± 19.17
Unweighted mean of weighted means (YBP)	690.61 ± 30.51	935.30 ± 4.01	1970.70 ± 16.31	720.86 ± 20.39
Weighted mean of the weighted means (YBP)	**672.46 ± 12.68**	**933.98 ± 13.37**	**1977.05 ± 14.59**	**720.16 ± 13.40**
χ^2 for weighted mean (2 degrees of freedom)	8.55	0.210	2.55	3.95
Significance level* (%)	1.39	90.1	28.0	13.9

* - The probability of obtaining, by chance, a scatter among the 3 laboratory weighted means as high as that observed, assuming the quoted random errors reflect all sources of variation.