[RADIOCARBON, VOL 32, NO. 3, 1990, P 329-334]

A SUGGESTED QUALITY ASSURANCE PROTOCOL FOR RADIOCARBON DATING LABORATORIES

AUSTIN LONG and ROBERT M KALIN

Environmental Radioisotope Center, Department of Geosciences
The University of Arizona, Tucson, Arizona 85712

ABSTRACT. The current intercomparison of data from ^{14}C laboratories reveals significant variability among liquid scintillation laboratories, suggesting that identical samples submitted to different laboratories may yield values that differ by much more than expected on a purely statistical basis. Erroneous dates (recently corrected) by a well-established ^{14}C laboratory give rise to further concern for quality ^{14}C data. Thus, it is incumbent on each laboratory to develop and implement a quality assurance and control (QA/QC) program in order to ensure accuracy of results and to alert lab personnel to problems.

Samples of pure materials (eg, benzene, cellulose) distributed by national or international standardizing groups are valuable, but are not representative of typical samples routinely run in most labs. Inevitably, ^{14}C personnel take special care with intercomparison samples and data that "outsiders" will be scrutinizing and comparing. Here, we reiterate Stuiver and Pearson's (1986) concept of laboratory error multiplier (K-value) and make the case for internally-generated QA/QC programs. We recommend that an ongoing, internal, self-test QA/QC protocol, to be designed and approved at the next ^{14}C conference, is the most practical and effective method of assuring quality of ^{14}C laboratory data. Each laboratory would then be responsible for determining its error multiplier factor by performing analyses on one or more homogeneous batches of wood chips, cellulose or calcite. Laboratories would update these data as they see fit and make this information available - in a standard format - to all who use their data.

INTRODUCTION

A significant number of users of ^{14}C data are losing the unquestioning confidence they once had in ^{14}C dates. They have heard that in a recent intercomparison study, some of the ^{14}C results of analyses on the same material were quite divergent. Evidence for this impression among the users comes from personal conversation, and even from an anonymous review of a National Science Foundation proposal (fortunately, a successful one). The ^{14}C dating community has an image problem. Even long-standing, well-established laboratories have not been immune to inaccuracy problems. This is a fixable problem. Some here have suggested fixes, and at least one of them is already underway. Here we also suggest a fix, which is not in competition with any of the other suggestions. In fact, it incorporates some elements of other plans.

QUALITY ASSURANCE IN ANALYTICAL CHEMISTRY

In cases of critically important analyses, both the laboratory and the consumer of the data make an effort to check the laboratory. The laboratory carries out elaborate steps to determine the analytical precision and accuracy of the product. In addition, unbeknownst to the laboratory, the wily consumer submits replicates and samples of known value as "unknowns" for analysis. The situation is somewhat like the Quantitative Analysis Lab exercises many of us had in college. The analyst is graded according to the quality of the data produced. At $200 to $500 per ^{14}C analysis, few radiocarbon data users have the resources and time to run their own independent checks of laboratory accuracy. And they should not have to.

The suggestion presented here is really nothing new, as it assimilates well-known and proven principles and procedures from analytical chemistry. These are collectively known as Quality Assurance (QA). Laboratories in the US that produce data that may become part of a lawsuit, such as our lab, must demonstrate their adherence to a quality assurance program.

A QA program is essentially a set of procedures that the lab personnel go through to convince themselves and others of the accuracy of their results, and establish the true precision of their analyses. A QA program is a continual rather than a one-shot process. In analytical chemistry, QA takes the form of repetitive analyses throughout the entire procedure of every chemical component which the laboratory reports analyses of, and in each different type of source material requiring distinctly different steps in the analyses. The analyst examines the time series for agreement with known values (accuracy), for dispersion of data (precision), and for trends with time and changes in procedure and personnel.

The present proposal differs from others we have heard so far in two respects:

1. All the responsibility and integrity, and nearly all the effort and expense lie within each laboratory.
2. The test samples would be similar to normal samples of geological and archaeological interest, *ie*, they would be samples that go through the laboratory system normally, as if they are unknowns, and with the fewest possible people in the lab knowing otherwise.

Special test samples in unusual chemical forms are not directly relevant to the question of precision and accuracy of the typical archaeological or geological sample:

1. Laboratory personnel know others will be scrutinizing their results and they take special care with them.
2. These abnormal samples will probably not go through the standard procedure of logging-in, handling, pretreatment and routine data checking. Thus they are not subject to "errors of the routine."

Special test programs establish an adversarial situation between those running the program and participating laboratories. Moreover, long time lags between the ^{14}C analysis and feedback of comparative results to the lab can lead to long delays in recognizing and fixing any problems in the labs.

The consumer of ^{14}C data needs to know, and should be provided, two types of information about the data:

1. How accurate is the ^{14}C analysis?
2. What does the "±" figure really mean?

All of us in the field of radiocarbon know that the ± figure is, by convention, a minimum estimate of uncertainty based on ideal counting statistics (Stuiver & Polach 1977). Many consumers of data are not aware of this, and correspondingly misuse the data statistically. It is useful to divide sample analytical precision into four levels:

Level 1. Derived from counting statistics alone

Level 2. Based on statistical analysis of repeat count rates of the same substance (such as CO_2 or benzene)

Level 3. Based on statistical analysis of count rates of samples repeatedly reprocessed through the entire procedure in the lab

Level 4. Based on statistical analysis of count rates of several samples re-sampled from the same stratigraphic level presumed to be an "instant" in time. (Actually, this is a field sampling, rather than laboratory element of the total precision question. Thus, the laboratory cannot evaluate Level 4 precision, but the consumer must consider this as a possible explanation for some inconsistencies.)

The ± figure, or precision, derived from each successive level is expected to be greater than the previous level.

An acceptable QA program must have the following attributes:

1. Evaluate the laboratory's accuracy for routine type samples.
2. Evaluate the Level 3 precision for a typical sample.

3. Allow laboratory directors to recognize problems immediately and begin remedial action.
4. Give the consumer confidence in the data a particular laboratory produced during a specific time interval.

THE PROPOSAL

1. Design, by committee, a recommended QA protocol, to include sample types, age ranges, frequency of repetition and data presentation.
2. Distribute large amounts of sample material to participating laboratories. Sample material would consist of at least two batches of fossil wood ca 1 and 2 half-lives old. The ^{14}C ages of both batches would be well established and known to all.
3. Participating laboratories would analyze these samples as unknowns at regular intervals, say monthly, and record the data graphically. Statistical analysis of these data would reveal bias, trends, sudden offsets and enable calculation of the Level 3 precision - total analytical precision - of ^{14}C analyses in each laboratory.
4. Laboratories would make these graphs available to anyone upon request. These laboratories would be authorized to include with data reports and publications a statement to the effect that "this laboratory adheres to the QA protocol recommended by . . . ". An error multiplier could also be on record so that their data would be statistically treated more properly.

Disadvantages of the Present Proposal

1. It is too easy for lab directors to "prune" the data. The success of this proposal depends on the scientific integrity of laboratory personnel.

Comment: All scientific endeavors depend on integrity at some point.

2. Some labs will consider this a waste of effort. They all run some standard.

Comment: "Oxalic acid only" has not worked for all labs. We still have an image problem. We need to make an extra collective effort to demonstrate accuracy.

3. Who will pay for preparation and distribution of QA samples?

Comment: We are attempting to set up an Association of Carbon-14 Labs (ACL). We would ask the Association only for mailing costs.

Advantages of the Present Proposal

1. It would allow for "instant" recognition by the laboratory director of analytical problems and an opportunity to remedy the situation quickly.
2. Realistic samples, which would go through normal laboratory channels, ideally would be unrecognized as QA samples by lab technicians.
3. It enables calculation of an operationally realistic figure of uncertainty which should be valid for statistical analysis of data.
4. This should be more acceptable to laboratories not willing to have "outsiders" knowing about problems before they do; it allows for ample "face-saving."
5. The protocol would provide the user with a quantitative, continuously updated evaluation of the quality of data emerging from each participating laboratory.
6. The present proposal would augment rather than replace other laboratory intercomparison studies.

Some Specific Examples

We offer some QA procedures that our ^{14}C lab undergoes to keep us sleeping well at night, that convince people for whom we generate ^{14}C dates that discrepancies are not due to laboratory problems, and that keep our lab on the list of U S Department of Energy approved laboratories for groundwater ^{14}C analysis. The laboratory technicians are, we suspect, unaware that the wood is a test sample.

1. *Two Creeks wood.* Figure 1 shows the results of a time sequence of preparations and analyses. Note from the error bars that we vary the size of the sample. The average of the ages is 11,902. The standard deviation (σ) calculated from the scatter of the analyses shown here is 157 years. These analyses also test our dilution techniques. Some of the analyses were on smaller portions of wood, and thus were more highly diluted with larger standard deviations. The average σ (based on counting statistics only) of the samples in Figure 1 is 149 years.

2. *Oxalic Acid I.* Figure 2 shows the results of a time sequence of separate preparations and analyses of ^{14}C and of $\delta^{13}C$ (NBS no. SRM 4990B). Here, OX I is run as an "unknown." The average in this series is 105.15 percent modern carbon (pMC). From the scatter of the individual runs, the σ is 0.53 pMC. The average σ (counting statistics only) of the runs listed here receiving normal counting times is 0.52 pMC.

Material of Known Age Available to All Laboratories

Wood of established age is available in large quantities. We would be able to distribute such wood to laboratory directors who assert that they will use it according to established protocol.

1. Two Creeks wood has been dated by several laboratories since WF Libby's original analyses. Several kilograms are presently languishing in our laboratory. An almost unlimited additional supply is available from this classic site.

2. Tree-ring dated fossil *Sequoia* from California, ~5000 - 6000 years old. Less of this is available.

3. Professor WG Mook, Groningen, has a large amount of cellulose that could fit into this QA program.

SUMMARY

1. A dark cloud is gathering over the radiocarbon community, which casts a shadow of doubt on the accuracy of some ^{14}C data.

2. It is incumbent on individual laboratories to dispel doubt by engaging in a formal QA program.

3. QA procedures are routine in most analytical chemistry laboratories, and these procedures are easily adaptable to natural ^{14}C analysis.

4. We propose that a small group of ^{14}C lab directors devise a protocol to be recommended to all laboratories that participate in the program.

5. Sufficient quantities of wood of well-established age could be made available to participating laboratories as International QA samples.

6. Participating laboratories could include a graph of analysis of these QA samples and their lab error multiplier as well as a statement regarding their adherence to "Approved QA Protocol" in data reports and publications.

7. This program would not only provide the user with needed information in addition to the ^{14}C date itself, but also improve general confidence in all ^{14}C data.

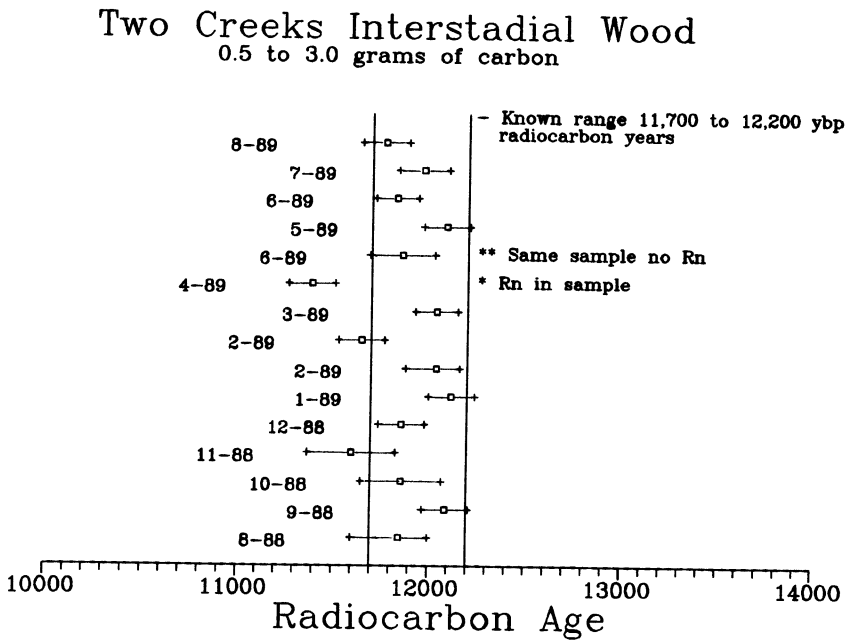

Fig 1. Two Creeks interstadial wood

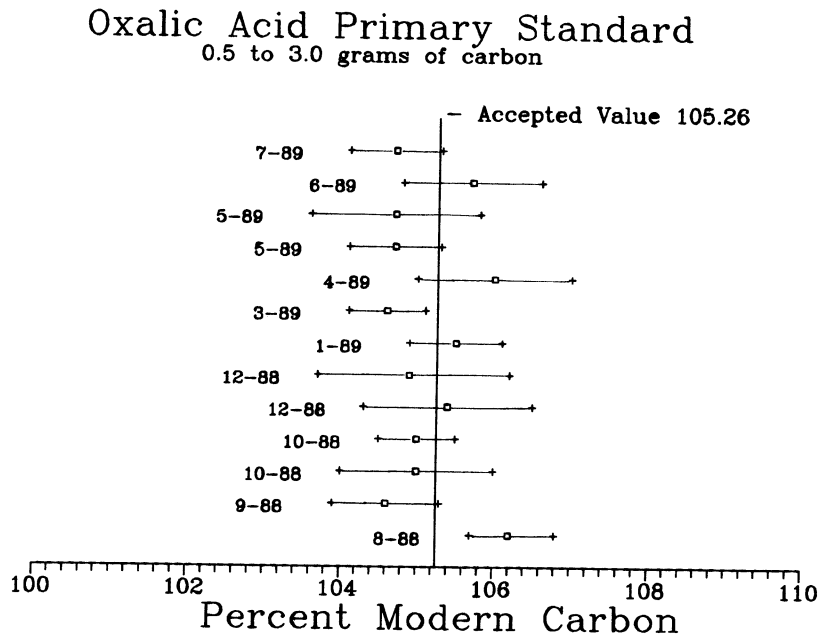

Fig 2. Oxalic acid primary standard

ACKNOWLEDGMENTS

These suggestions have evolved not only with experience, but also through extensive discussions with many colleagues on three continents. Special thanks to Wim Mook. NSF Grant no. BNS-8902140.

REFERENCES

Stuiver, M and Pearson, GW 1986 High-precision calibration of the radiocarbon time scale, AD 1950-500 BC. *In* Stuiver, M and Kra, RS, eds, Internatl ^{14}C conf, 12th, Proc. *Radiocarbon* 28(2B): 805-838.

Stuiver, M and Polach, HA 1977 Discussion: Reporting of ^{14}C Data. *Radiocarbon* 19(3) 355-363.