

ANNOUNCEMENT OF A FURTHER INTERNATIONAL INTERCOMPARISON EXERCISE

E. M. SCOTT¹, D. D. HARKNESS², B. F. MILLER², G. T. COOK³ and M. S. BAXTER⁴

ABSTRACT. Following recommendations of the Glasgow International Workshop on Intercomparison of Radiocarbon Laboratories (Scott, Long & Kra 1990), a further international intercomparison is planned. This new intercomparison is complementary to the existing IAEA intercalibration, and will make use of natural samples whose ages will be unknown to the participants. The study has been funded by the UK Research Councils (SERC and NERC), and samples will be free to all participants. We anticipate that this intercomparison will be ongoing, with distribution of samples in 1992, and presentation of the results at a later meeting. We present here details of the samples available and the time scale of the study. Briefly, we envisage that the new study will be more focused than the ICS (Scott *et al.* 1986), and will include natural samples in both pretreated and unpretreated forms.

INTRODUCTION

The need for a system of analytical quality control in applied radiocarbon dating is well recognized within the ¹⁴C community. Recently, a number of initiatives have been undertaken in this area; in this paper, we briefly describe two such initiatives and fully develop a third.

First, a suggested quality assurance protocol has been published (Long & Kalin 1990), which comprises a set of procedures that are the responsibility of each individual laboratory.

Second, a set of ¹⁴C reference materials has been prepared and distributed by the IAEA. The results of a major laboratory exercise to define precisely the new reference materials is the subject of other papers (Rozanski 1991; Rozanski *et al.* 1992). It is intended that laboratories should measure these reference materials at regular intervals, but that in addition, laboratories should develop their own laboratory standards which should be compared with the new reference materials.

The third tier of the initiative is the intention to hold a further international intercomparison, the description of which is the subject of this paper. Such intercomparisons provide the *only independent, objective and immediate assessment* of the maintenance of and/or improvement in analytical quality.

PREVIOUS INTERCOMPARISONS

A number of intercomparisons have been coordinated by the group at Glasgow University and SURRC; they can all be characterized by their use of typically natural samples whose results have not been known to laboratories in advance.

The first intercomparison organized by the group used tree-ring samples from a floating chronology. Laboratories received eight samples and were asked to return their results within a short time. The findings (ISG 1982) alluded to systematic laboratory bias and excess unexplained variability.

The second intercomparison has been the subject of a number of publications and was concluded by a workshop held in September 1989 (Scott, Long & Kra 1990). It was a much more ambitious

¹Department of Statistics, University of Glasgow, Glasgow, G12 8QW Scotland

²NERC ¹⁴C Laboratory, SURRC, East Kilbride, Glasgow G75 0QU Scotland

³Radiocarbon Laboratory, SURRC, East Kilbride, Glasgow G75 0QU Scotland

⁴IAEA Marine Laboratory, MC 98000 Monaco

exercise, involving 3 stages and 12 different samples (wood, shell and peat in both pretreated and untreated form). The summary of this intercomparison (Scott, Long & Kra 1990) again highlighted the possible existence of systematic laboratory bias and variability in excess of that quantified by the quoted errors.

On the basis of our experience in arranging the previous two international intercomparisons, we now discuss the issues as we perceive them and their proposed resolution in the third intercomparison (TIRI).

THIRD INTERNATIONAL RADIOCARBON INTERCOMPARISON (TIRI)

Simply stated, the aims of this further intercomparison are:

1. To function as the third arm of the quality assurance procedure
2. To provide an objective measure of the maintenance and improvement in laboratory procedures and, hence, analytical quality
3. To develop further our understanding of the validity and comparability of laboratory quoted errors
4. To assist in the development of a 'self-help' scheme for participating laboratories.

We address in turn a number of issues that are broadly defined as either a) scientific or b) administrative.

GENERAL DESCRIPTION OF TIRI

The second intercomparison involved a major effort by the ^{14}C community; the third intercomparison acknowledges the fact that, in two years, labs will have already completed the initial characterization of the IAEA ^{14}C reference materials. Thus, TIRI will take place, in the experimental phase, over a one-year period and will include 'natural' samples usually requiring full laboratory treatment, including pretreatment. Laboratories would typically be expected to select a minimum of six samples, selected from a range of materials on offer. However, a subset of samples would be compulsory to all laboratories to ensure a common base line in any analysis. As in the past, the precise activities of the samples will not be known to laboratories in advance.

SCIENTIFIC ISSUES

Identification of Systematic Biases: The Accuracy of a Laboratory

Since the intercomparison will incorporate material whose ages are known through other than radiometric means, typically supplied by dendro-dated tree rings, these samples will provide important information concerning any systematic laboratory offsets or biases. This clearly will not be possible for all samples, and for those whose activities are not known, the overall consensus values calculated in a manner similar to those for the IAEA reference materials will provide the base line against which the laboratory's results will be quantified.

Interpretation of Quoted Errors: The Precision of a Laboratory

1. The previous intercomparison incorporated unidentified duplicate samples to provide an objective measure of within-laboratory precision. The findings indicated that the lab-quoted errors, in most cases, adequately described the observed differences between duplicate samples, and allowed a quantification of the level of replicate error through evaluation of an error multiplier. TIRI will not further investigate this point in detail, as laboratory performance was, in general, satisfactory.

2. Both previous intercomparisons have indicated variability in excess of that quantified by the quoted errors (Scott *et al.* 1990), again quantified through the mechanism of an error multiplier, which included components defined by the calibration of results to an absolute standard (Ox1 and Ox2). Evidence of excess variation was also found in the IAEA results (Rozanski *et al.* 1992) which, in part, could be directly related to the modern standard used. Thus, TIRI should benefit from laboratory use of the new IAEA reference materials.

The definition of the laboratory error, when used within the framework of an intercomparison, should be based on the Level 3 definition of Long and Kalin (1990: 330) *i.e.*, "based on statistical analysis of count rates of samples repeatedly reprocessed through the entire procedure in the lab." The opportunity within an intercomparison of directly assessing the Level 3 error is, of course, strictly limited; we must depend on a clear evaluation of the errors quoted by the individual laboratory, and attempt to ensure that they are comparable. In addition, we can, as in the past, evaluate an error multiplier for each laboratory to ensure that, ultimately, after bias correction, the results, when considered as a whole, comprise a homogeneous group.

Comparison of Different Laboratory Types

In a number of previous studies, we have found evidence of different levels of performance dependent on the laboratory type (GPC, LSC or AMS). We envisage that such differences should be reduced with the introduction of the new quality assurance checks and the availability of the new IAEA reference materials. Nevertheless, comparability of laboratory performance will be investigated.

FURTHER DETAILS

Sample Material and Provenience

It is intended that samples should be of archaeological and geological interest, and reflect the broad spectrum of laboratory experience.

Sample material will include:

1. Marine turbidite and freshwater tufa, whole peat and humic acid samples, wood (some of which will have been dendro-dated)
2. Cellulose
3. Crystalline carbonate
4. A single-year grain sample.

In some instances, to ensure homogeneity, the sample will have undergone some preparation, *e.g.*, homogenization by slurring for the marine turbidite and extraction of humic acid for the peat. The wood will be identified tree-rings, a maximum of 40 years' growth.

The range of ^{14}C activities will span modern to background levels, with the majority of the samples concentrated in the range, 1 to 20 ka BP.

The homogeneity of the samples is of critical importance, and this will be tested for each sample before it is used in the intercomparison.

Reporting of Results

As in the past, a deadline for reporting results will be set. Following this, all participating labs will receive a copy of all the submitted results and a brief analysis. A full analysis will then be carried

out and a report provided to each laboratory within four months of the completion of the experimental phase.

Anonymity

The question of laboratory anonymity has proved important in the past. In previous intercomparisons, anonymity of laboratories has been preserved, with only a list of all participant laboratories being published. The recent IAEA reference work openly published identities of laboratories, but did permit laboratories to withhold their results. We intend that the consensus view of the community will determine this issue. We would encourage openness throughout. Nevertheless, we are aware of and concerned with the problems that can arise through a 'blacklisting' of laboratories, and thus, our intention will be to manage the intercomparison as a consultant for the participating laboratories.

Self Help

We have clearly stated that "a key component of the entire operation would be the help offered to participating laboratories . . ." (Mook 1990: 108). Clearly, with the identification of difficulties that a specific laboratory is suffering, it is important that *they be resolved in as short a time as possible and that they be shown to be resolved*. We would hope that, by maintaining an open policy, laboratories would, in most cases, be able to identify the appropriate labs/individuals to contact for advice. We would, of course, provide as much information as possible concerning the symptoms of any difficulties on the basis of the analytical data provided, and finally, would supply additional sample material if necessary to allow verification of the effectiveness of the remedial action.

Timetable and Progress

The timetable for the work described in preceding sections is as follows:

May 1991	Announcement of TIRI to coincide with ¹⁴ C conference Currently, laboratories are being contacted by letter concerning participation (over 90 positive responses have been received so far)
March 1992	Samples distributed to participating laboratories
April 1993	Experimental stage of intercomparison completed
Sept. 1993	Report circulated to participants
Aug. 1994	Concluding meeting, to coincide with the next ¹⁴ C conference.

CONCLUSIONS

This paper has discussed the role of a laboratory intercomparison within the framework of a quality assurance protocol, and has highlighted some of the key scientific and administrative issues. At this stage, it is intended as a consultative document, which we hope will stimulate discussion, and result in improved design for TIRI. We would be delighted to hear any comments on these or any other related issues.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the financial support of SERC and NERC.

REFERENCES

- International Study Group 1982 An inter-laboratory comparison of radiocarbon measurements in tree-rings. *Nature* 298: 619-623.
- Long, A. and Kalin, R. M. 1990 A suggested quality assurance protocol for radiocarbon dating laboratories. In Scott, E. M., Long, A. and Kra, R. S., eds., Proceedings of the International Workshop on Intercomparison of Radiocarbon Laboratories. *Radiocarbon* 32(3): 329-335.
- Mook, W. 1990 Special report from the Glasgow intercomparison workshop on quality control and assurance. *Radiocarbon* 32(1): 107-108.
- Rozanski, K. 1991 Consultants' group meeting on C-14 reference methods for radiocarbon laboratories, 18-20 February: 25 p.
- Rozanski, K., Stichler, W., Gonfiantini, R., Scott, E. M., Beukens, R. P., Kromer, B. and van der Plicht, J. 1992 The IAEA ^{14}C Intercomparison Exercise 1990. *Radiocarbon*, this issue.
- Scott, E. M., Aitchison, T. C., Harkness, D. D., Cook, G. T. and Baxter, M. S. 1990 An overview of all three stages of the international radiocarbon intercomparison. In Scott, E. M., Long, A. and Kra, R. S., eds., Proceedings of the International Workshop on Intercomparison of Radiocarbon Laboratories. *Radiocarbon* 32(3): 309-319.
- Scott, E. M., Baxter, M. S., Aitchison, T. C., Harkness, D. D. and Cook, G. T. 1986 Announcement of a new collaborative study for intercalibration of ^{14}C dating laboratories. *Radiocarbon* 28(1): 167-169.
- Scott, E. M., Long, A. and Kra, R. S., eds. 1990 Proceedings of the International Workshop on Intercomparison of Radiocarbon Laboratories. *Radiocarbon* 32(3): 253-397.