Future needs and requirements for AMS C-14 standards and reference materials
Scott, EM; Boaretto, E; Bryant, C; Cook, GT; Gulliksen, S; Harkness, DD; Heinemeier, J; McGee, E; Naysmith, P; Possnert, G

Published in: Nuclear Instruments & Methods in Physics Research Section B-Beam Interactions with Materials and Atoms

DOI: 10.1016/j.nimb.2004.04.074

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date: 2004

Link to publication in University of Groningen/UMCG research database

Citation for published version (APA):

Copyright
Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

Take-down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): http://www.rug.nl/research/portal. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.
Future needs and requirements for AMS $^{14}$C standards and reference materials

E. Marian Scott $^a$,*, Elisabetta Boaretto $^b$, Charlotte Bryant $^c$, Gordon T. Cook $^d$, Steinar Gulliksen $^e$, Doug D. Harkness $^c$, Jan Heinemeier $^f$, Edward McGee $^g$, Philip Naysmith $^d$, Goran Possnert $^h$, Hans van der Plicht $^i$, Mark van Strydonck $^j$

$^a$ Department of Statistics, University of Glasgow, Glasgow G12 8QW, Scotland, UK
$^b$ ESER Department, Weizmann Institute of Science, Rehovot, Israel
$^c$ NERC Radiocarbon Laboratory, Scottish Enterprise Technology Park, East Kilbride G75 0QF, Scotland, UK
$^d$ SUERC Radiocarbon Laboratory, Scottish Enterprise Technology Park, East Kilbride G75 0QF, Scotland, UK
$^e$ Radiocarbon laboratory, NTNU, N7491 Trondheim, Norway
$^f$ AMS Radiocarbon Laboratory, University of Aarhus, DK8000 Aarhus, Denmark
$^g$ Department of Experimental Physics, University College Dublin, Ireland
$^h$ Tandem Laboratory, University of Uppsala, Box 529, S-751 20 Uppsala, Sweden
$^i$ Center for Isotope Research, University of Groningen, Groningen, The Netherlands
$^j$ Royal Institute for Cultural Heritage, Brussels B-1000, Belgium

Abstract

$^{14}$C measurement uses a number of standards and reference materials with different properties. Historically the absolute calibration of $^{14}$C measurement was tied to 1890 wood, through the 'primary' standard of NBS-OxI (produced by the National Bureau of Standards, now NIST – National Institute of Standards and technology) subsequently replaced by NBS-OxII. These are both internationally calibrated and certified materials, whose $^{14}$C activities are known absolutely.

A second tier of materials, often called secondary standards or reference materials, and including internationally recognised materials such as ANU-sucrose (now also IAEA-C6), Chinese – sucrose and the IAEA C1–C6 series, augmented by additional oxalic acid samples are also used routinely. The activity of these materials has been estimated from large numbers of measurements made by many laboratories. Recently, further natural materials from the Third and Fourth International Radiocarbon Inter-comparisons (TIRI and FIRI) have been added to this list. The activities of these standards and reference materials span both the applied $^{14}$C age range and the chemical composition range of typical samples, but this is not achieved uniformly and there is a continuing need for reference materials for laboratory quality control and measurement-traceability purposes.

In this paper, we review the development of $^{14}$C standards and reference materials and consider the future requirements for such materials within the $^{14}$C AMS community.

© 2004 Elsevier B.V. All rights reserved.

PACS: 06.20.F; 93.85
Keywords: Metrology; Radiocarbon dating; Standards and reference materials

* Corresponding author. Tel.: +44-141-330-5125; fax: +44-141-330-4814.
E-mail address: marian@stats.gla.ac.uk (E.M. Scott).

0168-583X/S - see front matter © 2004 Elsevier B.V. All rights reserved.
1. Introduction

Reliable, precise and accurate $^{14}$C age measurements are essential for all areas of applied science where $^{14}$C is a key chronological tool. Such measurements also require traceability to international standards whose activities are known exactly by independent means and also to reference materials whose activities are estimated and typically accompanied by associated uncertainty statements. Within the $^{14}$C community, there has been an increasing realisation of the need for adequate reference materials. Long and Kalin [1] stressed that it was incumbent upon individual radiocarbon laboratories to engage in a formal programme of quality assurance (QA). Polach [2] noted that the opportunity for internal checking by individual laboratories involved in producing routine $^{14}$C measurements was hampered by a lack of suitable quality control (QC) and reference materials.

The internationally recognised primary standard is oxalic acid (NBS-OxI and subsequently NBS-OxII (now NIST)). While oxalic acid has all the physical and chemical attributes of a primary standard e.g. homogeneity, high purity, stability in storage, a constant and known gram molecular weight, etc., the quantitative and fractionation free recovery of its component carbon (initially as CO$_2$) has been problematic and the calibration of a material that is compositionally closer to the vast majority of samples submitted for radiocarbon analysis would be advantageous.

International efforts have resulted in the creation of a second tier of materials, so-called secondary standards or reference materials. These include materials such as ANU-sucrose, and the IAEA quality control reference series [3,4]. They have two main functions: for calibration, to demonstrate traceability and for quality control, to verify the performance of a laboratory.

For calibration, a reference material is often artificially produced so that its properties are known and well characterised with only a small uncertainty. For quality control purposes, a reference material is commonly a natural material so that it behaves as similarly as possible to the samples being measured. Most reference materials must therefore be certified on the basis of measurement by several laboratories, using different methods and using an independent calibration. Certification is only possible when agreement between laboratories performing the measurement can be demonstrated, usually in an inter-laboratory comparison.

The $^{14}$C dating community, over an extended period of time, has created a set of reference materials, resulting from a series of voluntary international inter-laboratory comparisons. The evolution of this voluntary programme is described and some of the issues in its operation are highlighted.

2. The historical development

Since the early days of applied $^{14}$C measurement it has been common practice for laboratories to exchange samples in attempts to improve and sustain analytical confidence. With time, this practice tended to give way gradually to a succession of more formal group inter-comparison exercises. Within the $^{14}$C community in just under 20 years, there have been a number of significant and very extensive inter-laboratory trials organised by individual laboratories and the International Atomic Energy Agency to the benefit of the $^{14}$C community (both labs and users) [3,5–11].

The comparisons have varied widely in terms of sample type and preparation, but have had as their primary goal the investigation of the comparability of results produced under possibly quite different laboratory protocols. In reaching this goal however, a number of these studies have also created reference materials. As methods and instrumentation have developed and new laboratories are formed, the reference materials created as a result of the inter-comparisons have been widely used for checking procedures and performance. Users have been reassured by the existence of regular comparisons that the laboratories are striving to ensure highest quality results while at the same time, the laboratories have been able to identify any systematic offsets and additional sources of variation. Indeed, in studies which have used representative samples requiring pre-treatment, chemical
synthesis and counting, it has been possible to identify the procedure in which problems have arisen and to quantify their relative contributions to the overall variation in the results. Thus participation in a laboratory inter-comparison has been seen to be a part of a formal QA programme and the resulting reference materials to form a community resource for the benefit of all.

3. Issues in the selection of reference materials

There are two key issues in the choice of \(^{14}\text{C}\) reference materials: whether the material should be natural or artificial and a demonstration of the material homogeneity. The design of the inter-laboratory trial for reference material certification is also the subject of some scrutiny.

3.1. Sample material

Reference materials can be either natural or artificial. For a natural material (such as peat, wood or shell), its \(^{14}\text{C}\) activity will not be known exactly, and there may be concerns about the homogeneity of the material. However, it will be representative of material that is routinely dated, therefore, the laboratory procedures used will be representative of those applied in routine measurement. For an artificial material, the activity and homogeneity may be known and easily demonstrated but the laboratory procedures used may be non-representative. In recent \(^{14}\text{C}\) dating inter-comparisons, predominantly natural materials have been used.

3.2. Homogeneity

A key question (particularly relevant when using natural samples) is the homogeneity of the material. Laboratory sample requirements, in terms of weight, may vary quite widely (as a result of differences in pre-treatment procedures and measurement technique in the different laboratories), thus it is necessary that the sample should be demonstrably homogeneous at the finest level required. Protocols and procedures for bulk material preparation and testing for homogeneity have been described and implemented [9,11,12].

3.3. Collaborative trial issues

Certification of reference materials based on measurements made by a number of laboratories is only possible when agreement between laboratories can be demonstrated.

For the \(^{14}\text{C}\) reference materials derived from IAEA, TIRI and FIRI, the certification procedures involved identification of any grossly anomalous observations followed by definition of a set of results that satisfy a homogeneity criterion (the criterion used defines a set of results to be consistent if they are less than two quoted errors away from a robust measure of sample activity). A weighted average of the set of consistent results is then used as the consensus value with its appropriate uncertainty [3].

When the trial is also used to assess individual laboratory performance, then design issues include the prescription (if any) of procedures that the laboratory should follow in detail, the format for reporting of results and the anonymity of participating laboratories.

Anonymity of laboratories is an issue that often concerns users. However, laboratory inter-comparisons only provide a spot check of operational performance at the time they are carried out and do not measure consistent performance over a period of time. Hence, they should not form the basis of a ‘league table of laboratory performance’ producing a ranking of laboratories. Inter-comparisons provide the participating laboratories with information that can be used to check and improve performance and in this way, they also provide indirect but important benefits to the user. Laboratories whose results are deemed problematic can be offered both assistance and additional samples so that any procedural difficulties can be identified and corrected. Recent \(^{14}\text{C}\) inter-comparisons have operated under this model.

4. Summary of past inter-comparisons

A summary of the inter-comparisons organised within the \(^{14}\text{C}\) dating community can be found in Scott et al. [13]. The first two collaborative trials [6,7,14] did not result in the creation of an archive...
of reference materials. However it is worth noting that following the International Collaborative Study (ICS) [7], the organisers concluded that some of the variations observed reflected the difficulties in maintaining suitable and sufficient laboratory standards and reference materials for calibration, and following this study, international efforts were made to extend the suite of reference materials that were available. It is also worth noting that by 1990, more than 50 $^{14}$C laboratories world-wide participated in the collaborative trial.

Following the ICS, six new potential reference materials were distributed in 1990 to over 130 laboratories for characterisation by the International Atomic Energy Agency [3]. This study was less concerned with laboratory performance but more with the suitability of the test materials and their future use. The materials had already undergone some homogeneity testing before distribution and ranged in activity from modern to background. Results from 69 laboratories were ultimately reported and formed the basis of the characterisation of the materials. Overall there was generally good agreement in the results, but a number of difficulties were subsequently identified with some of the materials. Problems associated with the storage of C-1 (Carrera marble) were identified and there was some evidence for contamination from the milling process of C-4 (Kauri wood) that resulted in up to 60% of the original results being excluded from the statistical evaluation. This highlights some of the practical difficulties in creating natural reference materials. The series C1–C6 has been recently supplemented with two additional materials, C7 and C8 which are both oxalic acids [4].

4.1. TIRI

TIRI (the Third International Radiocarbon Inter-comparison [8,10]) was begun in 1991, it again involved a large number of laboratories (over 70) and notably a substantial number of AMS laboratories. TIRI was designed to provide an independent assessment of laboratory performance, following the recently completed IAEA study and hence the materials were designed to test the full laboratory procedures. Sample processing included homogenisation by grinding and mixing. For some samples, (e.g. humic acid) chemical pre-treatment was applied before dispatch. All laboratories received a set of core samples. In the second stage, laboratories were able to select test samples from a list of materials that were of a more specialised nature and which might be seen as less routine. The optional samples included whalebone, whole peat, wood and travertine. The consensus values [8] for the samples were estimated using an approach similar to that taken in the IAEA study [3]. These were then used to provide an individual laboratory report on any laboratory bias and on laboratory precision.

The results from TIRI as in the other studies, pointed to variation in the results beyond that described by the quoted uncertainties. TIRI was not intended to explore the sources of the variation in the results, but it should be noted that there had been discussions at the TIRI workshop [8], concerning the homogeneity of the test samples, the issues of selection of small samples for AMS dating and the question of differing measured $^{14}$C contents depending on the chemical fraction dated. It is clear, that in any study using natural samples, some part of the extra variation must be due to the sampling of the bulk material. These issues are ones that have become increasingly important as the use of AMS has increased.

4.2. FIRI

The most recent laboratory inter-comparison was completed at the end of 2001. The set of core samples was distributed to over 120 laboratories during 1999 and by the deadline of December 2000, 92 sets of results had been received. These included results from more than 75% of operational AMS facilities [9,11,12] involved in $^{14}$C dating.

4.2.1. Sample selection, preparation and testing

Wood, peat, bone, marine carbonate (turbidite) and grain, together with specific fractions of samples such as the cellulose fraction of wood and the humic acid fraction of peat were identified as suitable materials. All bulk materials were prepared in a single batch and checked for
homogeneity by replicate analyses on eight randomly selected aliquots. The bulk samples were tested at different sub-sample sizes (reflecting one of the key differences between AMS and radiometric measurement) by two laboratories. The results of the homogeneity testing indicated that the samples could be considered to be homogeneous at the sample sizes tested and thus suitable as reference materials. It should be noted however that for the turbidite sample, a difference between the two testing laboratories was found and later demonstrated to be due to acid-leaching procedures used by one laboratory.

4.2.2. The inter-comparison

A set of seven core materials formed the basis of the inter-comparison. Three sets of duplicate samples were provided on a blind trial basis (Kauri wood, Belfast dendro-dated wood and Barley mash) to allow an assessment of laboratory precision resulting in a total of 10 samples to be analysed within a one-year period. Additional samples, which were not available in sufficient quantity for the main inter-comparison, or which were not routinely dated material (textiles, mammoth tusk, etc.) were offered as optional samples.

4.2.3. Consensus values for FIRI reference materials

Consensus values for the FIRI materials are shown in Table 1. The four dendro-dated wood samples included in the list of core samples were D and F (duplicates) from the Belfast master chronology and dendro-dated to 3200–3239 BC (14C age of 4495 BP); sample I (also from the Belfast master chronology) which has a dendro-date of 3299–3257 BC (14C age of 4471 BP) and sample H from the German oak chronology which was dendro-dated to 313–294 BC (14C age of 2215 BP). With respect to the dendro-dated samples, it can be observed that the consensus values and the average 14C ages from corresponding to the master calibration curve are such that the differences are all within the limits of the quoted errors. Thus, the consensus results are in agreement with the master chronology results and provide some confirmation of the validity of the consensus value estimation. A detailed analysis of the FIRI results appears in a special issue of Radiocarbon [12].

5. Conclusions

All of the studies cited above served several purposes. First they have provided valuable information to laboratories and hence to users. As a result, it is clear that such checks as TIRI and FIRI are and will continue to be necessary and that they must operate in addition to any internal laboratory procedures. Also it is not clear from the cited studies that the availability of existing reference materials has presented an immediate solution to the problem of laboratory comparability as might have been hoped. Second the studies have increased the numbers and types of reference materials (see Table 2). This is important, since by inclusion of a variety of reference materials in routine laboratory practice, the dating determinations can be better constrained.

<table>
<thead>
<tr>
<th>Sample description</th>
<th>Consensus value</th>
<th>Known age</th>
<th>Estimated precision</th>
</tr>
</thead>
<tbody>
<tr>
<td>A, B: Kauri wood</td>
<td>0.24 pMC(^a)</td>
<td>(95% CI (0.23–0.30))</td>
<td>–</td>
</tr>
<tr>
<td>C: Marine turbidite</td>
<td>18.176 yBP(^b)</td>
<td>–</td>
<td>10.5</td>
</tr>
<tr>
<td>D, F: Belfast dendro-dated</td>
<td>4508 yBP(^b)</td>
<td>3200–3239 BC (14C age 4495 BP)</td>
<td>3</td>
</tr>
<tr>
<td>E: Humic acid</td>
<td>11.780 yBP(^b)</td>
<td>–</td>
<td>7</td>
</tr>
<tr>
<td>G, J: Barley mash</td>
<td>110.7 pMC</td>
<td>–</td>
<td>0.04</td>
</tr>
<tr>
<td>H: Hohenheim wood</td>
<td>2232 yBP(^b)</td>
<td>313–294 BC (14C age 2215 BP)</td>
<td>5</td>
</tr>
<tr>
<td>I: Belfast cellulose</td>
<td>4485 yBP(^b)</td>
<td>3299–3257 BC (14C age 4471 BP)</td>
<td>5</td>
</tr>
</tbody>
</table>

\(^a\)Percent modern carbon.
\(^b\)Radiocarbon years before present where present is 1950.
Over the period in which these trials have been performed, there have been a number of significant developments in the $^{14}$C dating technique and as a result there is increasing user-driven pressure to date smaller (to the compound specific level) and older samples. There has been a substantial growth in the number of AMS facilities, and more radiometric laboratories are forming close collaborations with accelerator laboratories which has resulted in the growth in the number of laboratories preparing graphite targets. Perhaps, however, the most significant factor is that as the $^{14}$C community strives to measure smaller and smaller samples, the issue of sample homogeneity will become more and more important. Indeed the definition of a representative sample becomes critical. In some of the studies already completed in which AMS laboratories have participated, evidence of apparent sample in-homogeneity has been reported, undetectable at larger sample sizes.

With respect to the dating of old samples, close to the $^{14}$C limits of detection, there is a need for organic materials that are \textquoteleft\textquoteleft$^{14}$C-free\textquoteright\textquoteright. To achieve accurate and reproducible work, in particular with samples older than say three half-lives, it is essential to define the appropriate background signal using \textquoteleft\textquoteleft$^{14}$C-free\textquoteright\textquoteright material that has a chemical composition close to that of the sample. Ideally, such a reference material should also be subjected to an identical form of any pre-treatment that is applied to the raw sample.

Finally, it remains the case that continuing improvement in overall $^{14}$C quality relies on the identification and use of appropriate reference materials suitable for routine use.

Acknowledgements

This work was supported through Standards, Measurements and Testing, EC-SMT4-CT98-2265 and NERC GR9/03389.

References


<table>
<thead>
<tr>
<th>Table 2</th>
<th>Available reference materials from TIRI and FIRI</th>
</tr>
</thead>
<tbody>
<tr>
<td>TIRI-A</td>
<td>Barley mash</td>
</tr>
<tr>
<td>TIRI-F</td>
<td>Doublespar</td>
</tr>
<tr>
<td>TIRI-G</td>
<td>Fuglaness wood</td>
</tr>
<tr>
<td>TIRI-H</td>
<td>Ellanmore whole peat</td>
</tr>
<tr>
<td>TIRI-I</td>
<td>Travertine</td>
</tr>
<tr>
<td>TIRI-K</td>
<td>Turbidite</td>
</tr>
<tr>
<td>TIRI-L</td>
<td>Whalebone</td>
</tr>
<tr>
<td>FIRI-A and FIRI-B</td>
<td>Kauri wood</td>
</tr>
<tr>
<td>FIRI-C</td>
<td>Turbidite</td>
</tr>
<tr>
<td>FIRI-E</td>
<td>Humic acid</td>
</tr>
<tr>
<td>FIRI-G and FIRI-J</td>
<td>Barley mash</td>
</tr>
<tr>
<td>FIRI-H</td>
<td>Hohenheim dendro dated wood</td>
</tr>
<tr>
<td>FIRI-I</td>
<td>Belfast dendro dated cellulose</td>
</tr>
<tr>
<td>FIRI-K</td>
<td>Cambridge dendro dated wood and cellulose</td>
</tr>
</tbody>
</table>