Multiple radiocarbon dating of a single skeleton. Assessing issues of precision and accuracy in the Argaric Bronze Age.

(1) Vicente Lull, (2) Rafael Micó, (3) Cristina Rihuete-Herrada, (4) Roberto Risch, (5) Mark van Strydonck, (6) Mathieu Boudin

(1,2,3,4) Prehistory Department, UAB - Universitat Autònoma de Barcelona (Spain), (5,6) Royal Institute for Cultural Heritage in Brussels (Belgium)

(1) vicenc.lull@uab.cat, (2) rafael.mico@uab.cat, (3) cristina.rihuete@uab.cat, (4) robert.risch@uab.cat, (5) marbolleke@telenet.be, (6) mathieu.boudin@kikirpa.be

Abstract. The aim of this paper is to discuss issues of precision and accuracy of radiocarbon dates when testing a specific archaeological hypothesis. Our case study is framed in the Argaric “marriage versus descent” debate (Bronze Age of southeast Spain), where a high chronological resolution is needed in order to interpret the time gap between the deaths of the two adults of opposite sex usually found in double burials. We have dated the time of death of an Argaric male by means of six different skeletal samples. The results of this multiple dating of a single event support the precision of the method but raises questions over the limits of its accuracy when the human life span is on the same scale as the chronological resolution.

Múltiples dataciones por radiocarbono de un solo esqueleto. Evaluación de los problemas de precisión y precisión en la Edad del Bronce Argárico

Resumen. El objetivo de este artículo es discutir los temas de precisión y exactitud de las fechas de radiocarbono al probar una hipótesis arqueológica específica. Nuestro estudio de caso se enmarca en el debate argárico “matrimonio versus ascendencia” (Edad del Bronce en el sureste español), donde se necesita una alta resolución cronológica para interpretar la diferencia temporal entre las muertes de los dos adultos de sexo opuesto que suelen encontrarse en entierros dobles. Hemos datado el momento de la muerte de un varón argárico mediante seis muestras óseas diferentes. Los resultados de esta datación múltiple de un solo evento apoyan la precisión del método, pero plantean interrogantes sobre los límites de su exactitud cuando la duración de la vida humana está en la misma escala que la resolución cronológica.

1 Introduction: archaeological background and objectives

In the last decades, the increasing number of discoveries and analytical data from El Argar archaeological group offers a unique scenario for raising detailed hypotheses about the political and kinship relationships in an Early Bronze Age society. As a consequence, testing procedures need reliable and increasingly fine-grained analytical tools whose results are expected to settle debates, thus enhancing knowledge.

Argaric society has one of the most rich and diverse archaeological records in Late Prehistoric Europe (Fig.1). Dozens of permanent sites showing a well developed stone-built architecture and some of them reaching an urban status, shelter hundreds of single and double burials under living floors. Differences in grave goods assemblages among single burials were the source of a general consensus on the idea of a highly unequal distribution of wealth and power. In this sense, a hypothesis of State political organization supporting economic exploitation, social class division and inheritance rules for the transmission of property rights has been put forward [Lul86, Lul05].

1 For an overview of the archaeological definition of Argaric society in the context of the Early Bronze Age in Iberia and its current research agenda see [Lul11, Lul13a, Lul13b, Lul14, Lul16, Cam11, Ara15]
Double burials were excluded from statistical analysis because of the difficulties in allocating particular objects to specific individuals. On the other hand, double burials account for approximately 10% of the whole funerary Argaric record and the individuals buried together show different sex and age combinations, although adult/elderly male and female is the most common association. The bodies in these double tombs have been usually interpreted as a married couple in life, thereby endorsing the idea that Argaric kinship was built on monogamous nuclear families (e.g. [Inc870: 809, Sir87: 163-164; Chi58: 284, Sch12: 42]).

Our research on Argaric kinship includes DNA analysis still in progress and a programme of radiocarbon dating on samples from each one of the two adults found in double tombs. A test of 23 burials from nine Argaric sites showed that in most cases a.) the two radiocarbon dates from the same grave were statistically different, and b.) the time gap between the deaths of the alleged wife and husband accounted for at least two generations [Cas93, Lul00, Lul13b]. Therefore, it became clear that the hypothesis of conjugal bonds lacked support and that alternative views needed to be explored. We suggested that people buried together in Argaric double burials were linked by descent, thus ritually stressing the relevance of blood ties. It is expected that genomic data will contribute to further explore this issue.

The methodology involved in testing the ‘marital hypothesis’ relied upon the accuracy and precision of the radiocarbon dates for each of the skeletons buried in a double tomb. It is worth mentioning that only one single AMS-14C date was available for each skeleton, and that they were produced from anatomically diverse bone samples by different laboratories (OxA, KIA, Ua, Beta) since the late 1980’s. Moreover, other sources of uncertainty may have affected the reliability of the radiocarbon database used in our analysis: possible differences in sample collagen preparation among laboratories, differences in 14C fractionation according the type of bone and bone tissue, problems concerning sample quality (data on collagen content % and C% in collagen usually missing) or possible bias due to diets based on aquatic resources or influenced by climatic factors [Lul13b].

It is very difficult to overcome or rectify some of the shortcomings affecting the radiocarbon database (e.g. the lack of data on quality sample in analysis performed more than 20 years ago). Nevertheless, there may be other ways to assess the validity of the procedure used to test the archaeological hypothesis. This paper aims to address the reliability in the radiocarbon dating of human bone samples when testing specific archaeological questions. Our specific goal is to date the time of death of a single prehistoric human being by means of the radiocarbon analysis of different skeletal samples. If the results are statistically undifferentiated, the procedure (the radiocarbon measurement) should be deemed reliable enough and, regarding the double Argaric tombs debate, the questioning of the ‘marital hypothesis’ will be given an additional positive support. On the contrary, if the dates we get from the same human being are significantly different, this would mean low reliability and, in our case, that the large time gap between the pair of dates in most double tombs cannot be said with certainty to be a result of a real distance between deaths, but perhaps to a lack of precision of the measurement method.
2 Material and methods

The bone samples for this case-study were taken from the single skeleton found in tomb 67 of La Almoloya (AY67) (Pliego-Mula, Murcia province, Spain) [Lul15a, Lul15b], one of the best preserved Argaric sites which was occupied between 2200-1550 cal BCE (Fig. 2). The tomb is a large stone cist measuring 131 cm long and 62 cm wide (inner measurements), tightly closed until its excavation in August 2015 by means of calcareous slabs sealed with mortar. It had the remains of a 35-40 year old male lying on his back over a wooden board, with arms and legs tightly flexed to the right side and head rotated likewise, facing southwest (Figs. 3, 4). The scapula and radius of an immature sheep/goat found near the gluteal region are the only grave goods.

The skeleton found inside AY67 was selected for sampling because of the very good state of preservation upon its discovery. Bones are solid and devoid of the usual cracking, flaking and weathering effects of secular decay (Fig. 5). Taphonomic fractures are rare and very much localized. Root activity has left some characteristic imprints in compact surfaces but penetration into spongy tissue and medullar cavities is rare.

Six samples were selected in order to meet different sources of potential variability: a.) region: skull versus postcranium; b.) symmetry: upper versus lower extremities; c.) laterality: right versus left side; and d.) tissue type and remodelling rate: bones with high cortical mass (slow turnover) versus bones with high spongy mass (fast turnover) [Pri02]. All samples were sliced by means of a circular saw after a thorough anatomical examination of the skeleton, except for the petrous portion of the skull, which was found already detached from the rest of the temporal bone.

Sample processing, measurements and age calculation were conducted at the Royal Institute for Cultural Heritage in Brussels (Belgium). Collagen extraction was performed following the Longin method [Lon71]. A 1 % NaOH-wash was introduced between the demineralization and hydrolisation step for 15 minutes. Besides $^{14}$C, stable isotopes ($\delta^{13}$C and $\delta^{15}$N), %C, %N and C:N ratio were also analysed on the bone collagen.

Carbon and nitrogen stable isotope compositions were measured as the ratios of the heavy isotope to the light isotopes ($^{13}$C/$^{12}$C or $^{15}$N/$^{14}$N) and are reported in delta ($\delta$) notation as parts per thousand (%), where $\delta^{13}$C or $\delta^{15}$N = ($R_{\text{Sample}}/R_{\text{Standard}}$ - 1) $\times$ 1000, and R is $^{13}$C/$^{12}$C or $^{15}$N/$^{14}$N, relative to internationally defined standards for carbon (Vienna Pee Dee Belemnite, VPDB) and nitrogen (Ambient Inhalable Reservoir, AIR).

Figure 2 - Aerial view of La Almoloya with location of the tomb AY67.
Figure 3 - General view of the tomb before being opened (left) and after full skeletal exposure (right).

Figure 4 - The fully articulated skeleton in AY67. Although some minor displacements were observed, i.e. dispersal of some finger bones of the hand, opening of the rib cage under the weight of the left knee and shifting of the left arm, all of them account for a slow filling of the tomb once the body was already skeletonised.
Analyses were performed in duplicate on a Thermo Flash EA/HT elemental analyser, coupled to a Thermo DeltaV Advantage Isotope Ratio Mass Spectrometer via ConfloIV interface (ThermoFisher Scientific, Bremen, Germany). Standards used were IAEA-N1, IAEA-C6, and internally calibrated acetanilide. Analytical precision was 0.25‰ for both $\delta^{13}$C and $\delta^{15}$N based on multiple measurements of the standard acetanilide.

Graphitization of CO$_2$ for $^{14}$C dating was carried out using H$_2$ over a Fe catalyst. Targets were also prepared at the Royal Institute for Cultural Heritage in Brussels (Belgium) [Van90] and $^{14}$C concentrations were measured with Accelerated Mass Spectrometry (AMS) [Bou15]. $^{14}$C results are expressed in pMC (percentage modern carbon) and indicate the percent of modern (1950) carbon corrected for fractionation using the $\delta^{13}$C measurement. Calibrations of $^{14}$C dates were performed using OxCal 3 and the IntCal13 calibration curve data [Bro95, Bro01, Rei13].

The C:N ratio of the bone collagen samples was used to classify the collagen samples as uncontaminated or contaminated [DNi85, Amb90]. Samples providing results outside the 2.9 – 3.6 range were defined as contaminated, since a higher C:N is the result of introduction of exogenous carbon-containing compounds (i.e. contamination).

Carbon and nitrogen concentrations in bone gelatin in relation to the bulk weight were also determined and will be referred to as weight percentage of carbon and nitrogen (%C and %N). These two quality indicators provide information on protein degradation. A higher %C than the average %C of modern bone collagen may indicate contamination. Ambrose (1990) cites a collagen weight %C and %N range for modern bone and tooth between 15.3% and 47% and 5.5 to 17.3% for C and N, respectively.

The amount of extracted collagen (gelatin) in relation to the weight of the whole bone sample will be referred to as collagen weight proportion in percent (% collagen). The % collagen is a quality indicator for collagen preservation. The threshold used in our laboratory is 2% and bones containing less than 2% collagen are considered poorly preserved.

3 Results and discussion

Five out of the six samples gave valid results. They are listed in table 1 and plotted in fig. 6. The bone sample from the skull (left petrous portion) did not contain any collagen.

The C:N ratio of all the samples (see Table 1) fall within the C:N range of uncontaminated bone collagen. The %C, %N and collagen yield (see Table 1) indicate in all cases well-preserved bone collagen.

The stable isotope results of all the samples, taking into account the analytical precision of 0.25‰, are in perfect agreement and demonstrate the validity of the $^{14}$C dates. The stable isotope values indicate a terrestrial diet and exclude marine or freshwater consumption. Consequently, no reservoir effect is present on the $^{14}$C ages.
Table 1 - Radiocarbon dates from the male buried in La Almoloya tomb 67 (calibration curve: IntCal2013 –[Rei13]).

<table>
<thead>
<tr>
<th>Lab. code</th>
<th>BP Cal</th>
<th>Calibration (cal BC)</th>
<th>Sample description &amp; weight</th>
<th>%N</th>
<th>%C</th>
<th>δ¹⁵N‰ corr</th>
<th>δ¹³C‰ corr</th>
<th>C:N</th>
<th>% collagen</th>
</tr>
</thead>
<tbody>
<tr>
<td>RICH-22901</td>
<td>3545±31</td>
<td>68.2% prob. 1940 (49.8%) 1870 1850 (11.1%) 1820 1800 (7.3%) 1780 95.4% prob. 1990 (95.4%) 1760</td>
<td>Left humerus: anterior distal shaft, cortical fragment (3.8 g)</td>
<td>12.24</td>
<td>34.1</td>
<td>8.8</td>
<td>-19.4</td>
<td>3.2</td>
<td>2.7</td>
</tr>
<tr>
<td>RICH-22902</td>
<td>3525±32</td>
<td>68.2% prob. 1910 (23.7%) 1860 1850 (44.5%) 1770 95.4% prob. 1940 (95.4%) 1750</td>
<td>Right humerus: anterior distal shaft, cortical fragment (2.5 g)</td>
<td>9.45</td>
<td>26.1</td>
<td>8.6</td>
<td>-19.4</td>
<td>3.2</td>
<td>4.2</td>
</tr>
<tr>
<td>RICH-22903</td>
<td>3522±30</td>
<td>68.2% prob. 1900 (19.4%) 1860 1850 (48.8%) 1770 95.4% prob. 1930 (95.4%) 1750</td>
<td>3rd left rib, body fragment (1.7 g)</td>
<td>12.23</td>
<td>33.9</td>
<td>8.5</td>
<td>-19.4</td>
<td>3.2</td>
<td>10.1</td>
</tr>
<tr>
<td>RICH-22907</td>
<td>3512±30</td>
<td>68.2% prob. 1990 (14.3%) 1860 1850 (53.9%) 1770 95.4% prob. 1920 (95.4%) 1740</td>
<td>Complete proximal hand phalanx (2nd or 3rd finger; unknown laterality) (3.9 g)</td>
<td>13.53</td>
<td>37.5</td>
<td>8.5</td>
<td>-19.3</td>
<td>3.2</td>
<td>6.2</td>
</tr>
<tr>
<td>RICH-22905</td>
<td>3497±32</td>
<td>68.2% prob. 1990 (68.2%) 1770 95.4% prob. 1910 (95.9%) 1740 1710 (1.5%) 1690</td>
<td>Left femur: posterior middle shaft, cortical (4.5 g)</td>
<td>12.80</td>
<td>35.2</td>
<td>8.5</td>
<td>-19.3</td>
<td>3.2</td>
<td>4.2</td>
</tr>
</tbody>
</table>

Pooled mean of the former five dates

3520±14

δ¹⁵N‰ corr | δ¹³C‰ corr | C:N | % collagen |

12.80 | 35.2 | 8.5 | -19.3 | 3.2 | 4.2 |

\( X^2\)-test: df=4 T=1.3 (5% 9.5)
The difference of only three radiocarbon years between the cortical fragment of the left humerus (RICH-22901) and the third left rib fragment (RICH-22903), which was richer in spongy bone, shows that tissue type does not affect the results. On the other hand, a larger difference between the left and the right humerus (RICH-22902) (20 radiocarbon years) suggests that sampling the same bone does not necessarily provide much closer results. The same can be said in relation to laterality and cortical tissue, since the largest distance between any pair of dates is the one between the left femur (RICH-22905) and the left humerus (48 radiocarbon years), both of them equally rich in dense cortical tissue.

Nevertheless, the comparison between the oldest (RICH-22901) and the youngest (RICH-22905) values of the series shows that both are statistically the same (T=1.160; Chi²: 3.84, df: 1; p <0.05). This means that the difference of 48 radiocarbon years between their central values does not have a statistical relevance and that all five results can be regarded as a single date. Therefore, a first conclusion to be drawn is that radiocarbon dating of human bone samples may be considered a rather precise analytical tool or, in other words, that the measurement of a single physical event (the death of a human being) has produced a set of values about which it can be said they are statistically the same. Moreover, this supports the research conclusions on Argaric double burials, in the sense that we can feel confident that a statistically significant difference between any pair of dates from the two skeletons buried in the same grave reflects a truly large temporal gap between both deaths.

What are the implications of the current degree of accuracy in radiocarbon dating? Are we able to obtain dates accurate or not. In this sense, precision expresses reproducibility and, consequently, we cannot evaluate the precision of a single measurement.

Accuracy, on the other hand, is how close a measurement is to its reference (“true”) value. Therefore, accuracy cannot be expressed as in precision (i.e. by means of measures of statistical dispersion), since accuracy increases insofar the measurement error decreases.

Our test has shown that radiocarbon dating with modern standards is a precise method because the results of different measurements agree closely with each other. Yet, how are we going to estimate its accuracy when the reference value (i.e. time of death of that specific human being) is unknown? We should rather focus on the level of uncertainty resulting from statistical analysis and its meaning in historical terms. Uncertainty in radiocarbon dating depends, in the first place, on the width of the standard deviation of each measurement (nowadays mostly between ±20 and ±35 radiocarbon years), and on the wiggling of the segment of the calibration curve intersected by a particular date.

In this case-study, the calibration of the pooled mean from all five dates gives probability ranges of 110 (1 sigma) and 140 (2 sigma) calendar years (Table 1). In the context of debating the ‘marriage’ versus the ‘descent’ hypothesis on Argaric double burials, can we reasonable expect to achieve any successful testing having such large uncertainty ranges?

4 Conclusion

Two physical events (as, for example, human deaths) several decades apart might produce (statistically) identical radiocarbon dates, but two statistically different radiocarbon dates most probably do not refer to two physical events taking place close in time.

This statement, which is all the more relevant when testing archaeological hypothesis, depends highly on the temporal gap between a series of physical events. In the case of double Argaric burials, additional support to the critique of the ‘marriage hypothesis’ has been given by means of strengthening the reliability of the conclusions put forward by Lull et al. ([Lul13b]; that is, that in most cases the second inhumation took place at least several decades after the first one. Indirectly, it favours the ‘descent hypothesis’, although its testing needs further support (i.e. positive results from DNA analysis).
Radiocarbon dating is a precise tool for testing archaeological hypothesis. Hopefully, higher accuracy will allow new advances for addressing actual daring hypothesis about past social relationships.

Acknowledgments

The results presented in this text have been produced by investigations funded by the Spanish Ministry of Economy and Competitiveness (project HAR2014-53860-P). We are also grateful for the support of CEFU, S.A. and Pliego town Council during the archaeological excavations carried out at La Almoloya, as well as the cooperation of the scientific and technical staff of the ‘Bastida Project’ and the Research Group for Mediterranean Social Archaeoecology (ASOME – 2014SGR919) of the Autonomous University of Barcelona (UAB).

Bibliography


