

The Bible and Radiocarbon Dating

Archaeology, Text and Science

edited by

Thomas E. Levy

&

Thomas Higham

Equinox Publishing Ltd

equinox

London Oakville

2005

14 Quality Control of Groningen ^{14}C Results from Tel Rehov

Repeatability and intercomparison of Proportional Gas Counting and AMS

Johannes van der Plicht and Hendrik J. Bruins

Abstract

Stratified radiocarbon dates provide a scientific chronological framework independent of cultural assessments. In both Groningen ^{14}C labs (conventional and Accelerator Mass Spectrometry [AMS]), a total of 64 radiocarbon dates were measured from Tel Rehov, derived from 21 Iron Age loci. This is the largest Iron Age series available at present for any site in the Near East. We present, evaluate, and discuss in this article our methodology in terms of quality assurance, reliability and reproducibility.

Introduction

Radiocarbon dating plays a key role in (pre)historic research, because it provides a scientific yardstick (quite often the only objective) for the measurement of time. This yardstick is independent of cultural deliberations and enables chronological comparisons, for example, of different areas at an excavation site, or between sites and regions. This is essential for proper interpretation of archaeological layers and association with other data (van der Plicht and Bruins 2001).

Our ^{14}C Iron Age chronology established for Tel Rehov, Israel, unambiguously favors a 'high chronology' (Bruins, van der Plicht, and Mazar 2003a, 2003b). However, our findings are contested by Finkelstein and Piasezky (2003; Chapter 16, this volume). It is of crucial importance to be aware of perceived or real quality problems of dating results. The present Iron Age chronology discussion takes place at the limits of resolution of the ^{14}C method, as issues need to be resolved with a temporal resolution well within a century. Here even small errors may have important consequences. Mistakes can be made by the ^{14}C laboratory (methodology and accuracy), and/or in the field (sampling and association).

In this paper we discuss the ^{14}C dates from Tel Rehov in terms of quality control, as measured by the two Groningen radiocarbon laboratories: conventional by means of Proportional Gas Counting (PGC) and AMS. This is necessary in order to evaluate the coherence and robustness of our radiocarbon series, composed of 64 dates, based on two separate measurement systems (for all these dates, refer to the tables in Mazar *et al.* [Chapter 13, this volume]). Our dates are contested by proponents of the 'low chronology' theory. They refer to other ^{14}C dates, which are comparatively younger than the Groningen series, as far as Tel Rehov is concerned (Bruins, van der Plicht, and Mazar 2003b, 2004).

The ^{14}C Dating Method

Radiocarbon (^{14}C) is a natural isotope of the element Carbon, which is produced by cosmic radiation and occurs in minute concentrations in living organisms. Since this isotope is radioactive (with a half-life of 5730 years) the ^{14}C concentration gradually declines after the death of the organism. Thus, by measuring the remaining ^{14}C content in (pre)historic samples one can date these samples, or more precisely, calculate the moment of death of the organism.

Although the principle of ^{14}C dating is based on this straightforward model, in practice the theory is complicated by many factors. First, the natural ^{14}C concentration has not been constant through time; the exact value of the half-life to be used for ^{14}C age calculations is not transparent; furthermore, isotopic effects in natural processes change the ^{14}C content for various sample materials, thereby changing the age. Second, measuring the ^{14}C radioactivity in natural concentrations is not simple, and requires specialized laboratories and measuring procedures. Here we summarize briefly the issues relevant for this paper. For more detailed accounts we refer to the specialized literature.

The theoretical complications are solved by a number of conventions (e.g. Mook and Waterbolk 1985). The ^{14}C measurements are reported in a specially defined unit BP. This definition comprises: (a) the ^{14}C radioactivity is measured relative to an international standard (i.e. oxalic acid with a ^{14}C activity corresponding to 1950 AD); (b) the ^{14}C age is calculated using the original half-life value of 5568 years; (c) the ^{14}C age calculation includes correction for isotopic fractionation using the ^{13}C content of the sample (to the standard value of $\delta^{13}\text{C} = -25\text{‰}$).

This convention is valid for all measuring techniques used for ^{14}C dating. Thus, the ^{14}C timescale is *defined*; a ^{14}C chronology (in BP) is different (and varying over time) from a calendar chronology (in BC/AD). The relation between the two timescales is established by calibration, which is the process of translating ^{14}C dates into calendar ages. Calibration curves are established by high precision ^{14}C measurements of wood samples dated absolutely by dendrochronology. The present calibration curve recommended by the ^{14}C community is INTCAL98 (Stuiver *et al.* 1998) to be replaced by INTCAL04 (Reimer *et al.* 2004). The part of the calibration curve relevant for this volume, 1500–500 BCE, is shown in Figure 14.1.

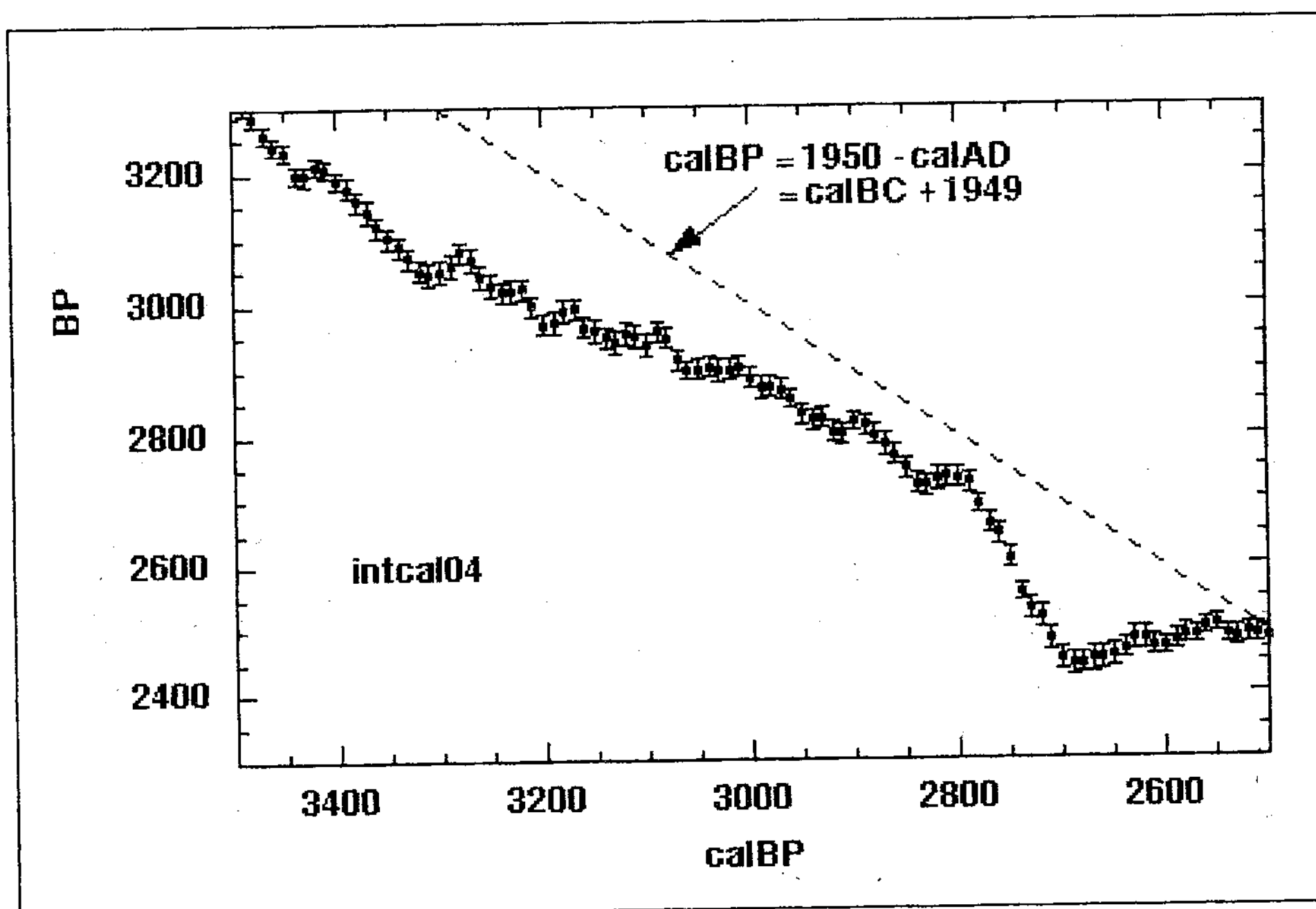


Figure 14.1. Radiocarbon calibration curve (Reimer *et al.* 2004) for the time range 1500–500 BCE.

Calibrated age range distributions for individual or stratified series of ^{14}C dates can be obtained using computer programs like OxCal (Bronk Ramsey 1995).

The measuring process of the ^{14}C content of archaeological samples such as bone, charcoal, wood, and so on, can be viewed from two perspectives: measuring technique and sample preparation. For ^{14}C , three different measuring techniques have been developed—PGC, Liquid Scintillation Spectrometry (LSC), and AMS. The first two, PGC and LSC, are called ‘conventional techniques’ and are based on radiometry; they require ca. 1 gram of carbon (e.g. Kromer and Münnich 1992; Theodorsson 1996). Larger sample amounts may result in a more precise measurement with a smaller standard deviation. The technique of AMS is based on mass spectrometry, for which mg size C is sufficient (e.g. Bayliss, McCormac, and van der Plicht 2004; Tuniz *et al.* 1998). The radio-carbon laboratory in Groningen houses both a conventional (PGC-based) and an AMS laboratory, with both facilities operating independently.

Sample preparation follows similar procedures for all three measuring techniques (Mook and Streurman 1983). The general rules are that contaminants have to be removed (physically and chemically) and that a reliable datable fraction has to be isolated. The latter is then combusted into CO_2 gas, which needs to be purified. For PGC, the $^{14}\text{CO}_2$ is counted in terms of radiometry; for AMS, the CO_2 needs to be transferred into graphite. Stable isotope mass spectrometers are used to measure the $\delta^{13}\text{C}$ content of the same CO_2 gas. Apart from fractionation information, the $\delta^{13}\text{C}$ is a measure for quality of the sample material—as well as the carbon content.

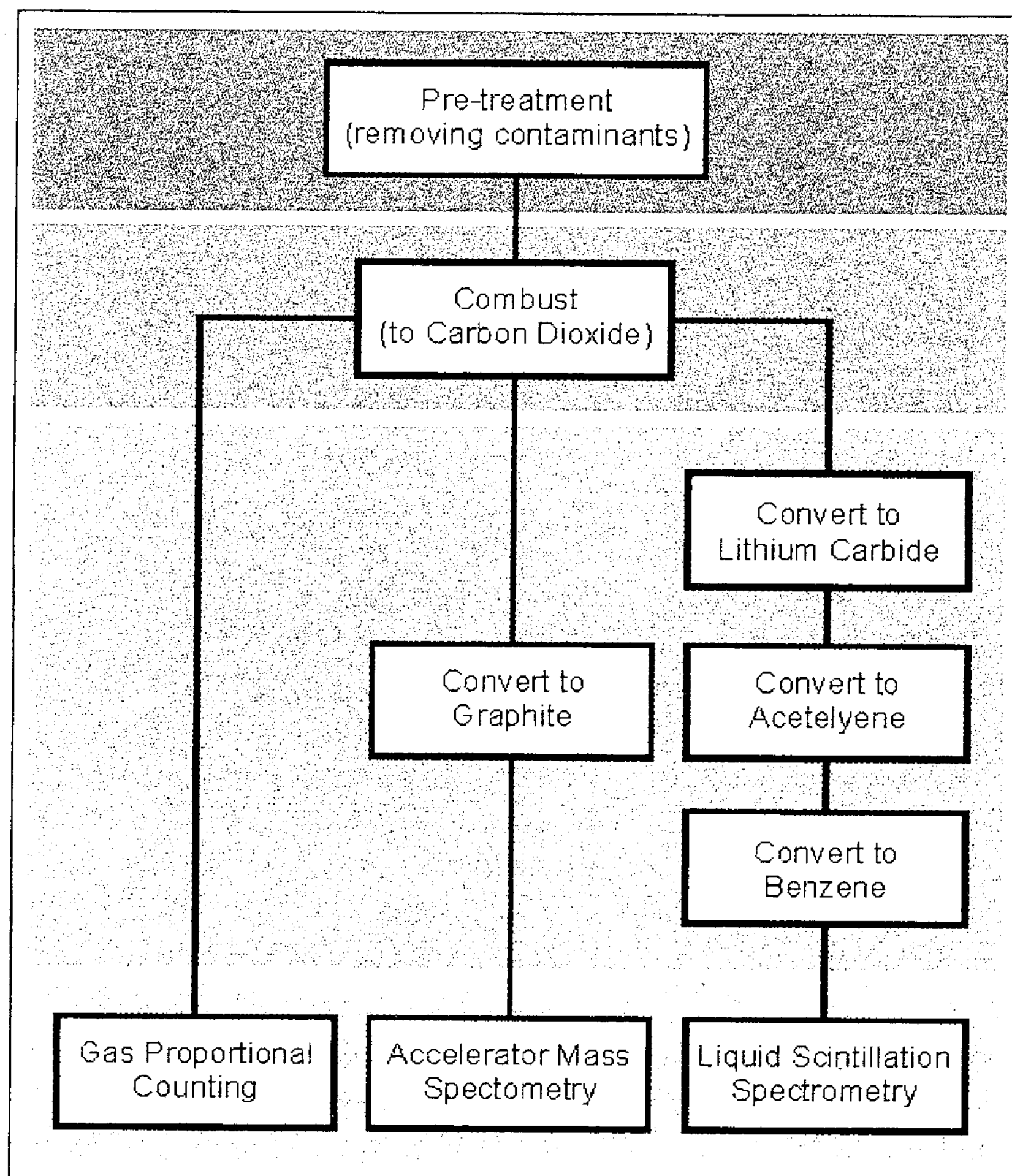


Figure 14.2. Schematic illustration for measuring archaeological samples (Bayliss, McCormac, and van der Plicht 2004; Mook and Streurman 1983): (a) overview of methods for measurements of ^{14}C ; (b) method of pre-treatment for an archaeological charcoal sample.

The pre-treatment for an archaeological charcoal sample is shown as an example in Figure 14.2. Only the ^{14}C that was part of the organism when it died should be measured. Therefore the first task is to remove any foreign carbon that has entered the sample material since that time. Such contamination comes mainly from the burial environment. We mention the two most common contaminants for charcoal samples: rootlets (usually from higher strata and thus younger) may be present in the collected material; soil carbonate (containing fossil carbon, thus older than the sample). We refer to Mook and Streurman (1983) for more details.

Quality Control Issues

Radiocarbon (^{14}C) is the most common radiometric dating tool applied in disciplines such as archaeology, geosciences, and environmental research. Stringent quality control is required to build up reliable ^{14}C chronologies, in particular for the (pre)historical periods in the Levant. Important aspects of quality control involve regular laboratory intercomparisons, multiple analyses of selected samples, sample material selection, archaeological association, and sample size (i.e. conventional versus AMS as measuring technique). These aspects will be discussed below.

General Aspects

True point dates cannot be achieved with ^{14}C dating, as there will always be a standard deviation. Both equipment resolution and measurement stability, as well as the random nature of radioactive decay (Mook and Waterbolk 1985: 10) 'causes the results of repeated measurements to spread around a "true" value. The possible discrepancy between a measured value and the "true" value is indicated by the standard deviation (σ)'. Hence one measurement of a sample may result in a date that is very close to the real date, in terms of the mid-point Gaussian value, but may also be older or younger, depending on the value of the standard deviation. Concerning radiocarbon research in Near Eastern historical archaeological periods in which an accuracy and precision is required within a 100 year BP range, multiple measurements of the same sample were suggested (van der Plicht and Bruins 2001). Such multiple measurements will theoretically result in an average date that is both more accurate and precise than can be achieved with single measurements, provided that the ^{14}C laboratory does not have a systematic bias towards older or younger dates. The issue of multiple measurements is further discussed below.

The quality of the BP date—before calibration—always forms the basis for every radiocarbon age determination. It must be realized that a ^{14}C date *does* provide a very important universal physical measurement of time, independent of cultural–historical viewpoints and associative reasoning.

Sample selection is a critical component in the ^{14}C dating process. The layers from which archaeological or geological samples are taken during excavations have not always remained static and may have been affected by different kinds of post-depositional processes. Perturbation by plants, animals or human activities (e.g. digging) may cause migration or contamination of carbon in samples used for ^{14}C dating.

Another key question is the relationship between the age of the sample and the archaeological or historical question addressed: 'How is the ^{14}C event related to the human event to be dated?' (van Strydonck *et al.* 1999). A well known problem in this respect is the so-called 'old wood effect'. Wood used (or re-used) to construct a building may have a ^{14}C date that differs from the human construction event by several centuries, depending on the age of the wood. It must be emphasized that the ^{14}C date of the wood in such a case is *not* a measurement mistake. Rather the age of the wood sample is older than the age of the archaeological layer or building in which it was found.

Another important matter related to sample selection is the respective choice of 'conventional dating versus AMS'. There can be a temptation to collect and submit all isolated seeds and tiny flecks of charcoal. The dating of such isolated samples by AMS should be discouraged, if larger samples (seed or charcoal clusters) are present in the same layer. If sufficient material is available, samples can be dated more cheaply and often more accurately by conventional means. The possibility of dating erratic post-depositional influences is considerable when isolated small fragments of charcoal or seeds are used, which are liable to movement by faunal or human digging activity. Such tiny samples have to be derived from a clearly defined context or association to justify dating. Lanting and van der Plicht (1994) presented a detailed discussion about these issues, including examples. It is a 'myth' that AMS is better than conventional radiocarbon dating: standard deviations are usually not smaller.

Time-width effects represented by a sample have to be considered. Bulk samples of peat layers, for example, are centimeters thick for conventional ^{14}C analysis. Such a sample comprises many years of sedimentation or growth. Isolated seeds, macrofossils, and grains represent single-year samples and are typical AMS material, due to their small sample size, but the stratigraphic context must be clear, as noted above.

The correct calibration procedure of ^{14}C dates from multi-year or single-year samples needs to be carefully contemplated. Smoothed curves are recommended for multi-year samples, while single-year samples ought to be calibrated with the most detailed calibration curve available (Mook and Waterbolk 1985).

The 975–955 BCE wiggle, so important in the Iron Age (Bruins, van der Plicht, and Mazar 2003a, 2003b; Mazar *et al.* [Chapter 13, this volume]) is based on dendrochronological data through radiocarbon measurements of groups of 10 tree rings (decadal) or 20 tree rings (bidecadal) (Stuiver *et al.* 1998). The radiocarbon measurements of Tel Rehov Stratum VI are based on short-lived samples (seeds, bone) with a time span of a few years or even less (growing season). Past ^{14}C variations in the atmosphere are not captured in such detail by the calibration curve itself, due to the comparatively 'coarse' decadal and bidecadal measurements. Therefore, it might be important to make a special primary study of this 975–955 BCE wiggle through ^{14}C measurements of each individual tree ring in the available dendrochronological series for the period 980–940 BCE or even for the entire 10th century BCE. Such detailed measurements have been carried out for older Bronze Age periods (Vogel and van der Plicht 1993).

Intercomparison

Intercomparison is a major part of quality assurance. In an intercomparison, the same samples are dated by different laboratories. This may involve either samples of known age or blind samples. Thus a laboratory can 'check' its performance—in particular, the sample (pre)treatment and ^{14}C measurement procedures. Intercomparison is a well recognized issue in the ^{14}C community, and various exercises form a continuing process. The latest large scale intercomparison is FIRI (Fourth International Radiocarbon Intercomparison) in which 84 laboratories participated worldwide. Several publications are generated by this program (Boaretto *et al.* 2003; Scott *et al.* 2004), and the final report is a special publication of the *Radiocarbon* journal (Scott [ed.] 2003).

This FIRI had the following aims:

1. Evaluation of the comparability of routine analysis of both AMS and conventional laboratories.
2. Quantification of the extent of and the sources for any variation.
3. Investigation of the effects of sample size, precision, and pre-treatment on the results (Scott [ed.] 2003).