Circaea

Circaea is the Journal (formerly Bulletin) of the Association for Environmental Archaeology (AEA) and—as from Volume 4—it is published twice a year. It contains short articles and reviews as well as more substantial papers and notices of forthcoming publications.

The Newsletter of the Association, produced four times a year, carries news about conferences and the business of the Association. It is edited by Vanessa Straker, to whom copy should be sent (c/o Department of Geography, University of Bristol, Bristol BS8 1SS).

Editorial policy for Circaea is to include material of a controversial nature where important issues are involved. Although a high standard will be required in scientific contributions, the Editors will be happy to consider material the importance or relevance of which might not be apparent to the editors of scientific and archaeological journals—for example, papers which consider in detail methodological problems such as the identification of difficult bioarchaeological remains.

Circaea is edited and assembled by Allan Hall and Harry Kenward at the University of York, England. It is distributed free to members of the AEA. Annual subscriptions to the Association are currently £12 for ordinary individual members, £6 for students and unwaged, and £18 for institutions. The institutional subscription to Circaea is also £18.

At present, copyright resides with individual authors and with Circaea, on the understanding that Circaea may reprint material without consulting the author, and that the author may reproduce their own articles without permission. This is intended to circumvent the complexities of photocopying restrictions. Circaea waives restrictions on multiple copying for academic purposes by anyone but copies so made should not be sold.

Circaea is published by the Association for Environmental Archaeology, c/o Environmental Archaeology Unit, University of York, Heslington, York YO1 5DD, U.K. Enquiries concerning membership of the AEA should be sent to the Membership Secretary, Association for Environmental Archaeology, c/o University Museum, Parks Road, Oxford OX1 3PW, U.K. Notes for contributors may be found on the inside back page of this issue.

Front cover: Based on a distribution map generated by Philippa Tomlinson and the University of York mainframe computer from the ARCD (see pp. 1-30)
Design and implementation of a relational database for archaeobotanical records from Great Britain and Ireland

Philippa Tomlinson

Environmental Archaeology Unit, University of York, York YO1 5DD, U.K.

Summary

The development of a comprehensive computer database of published records of macrofossil plant remains identified from archaeological sites in Great Britain and Ireland is described. The Archaeobotanical Computer Database (ABCD) has been established in a relational database management system using Structured Query Language (SQL). An introduction is provided, both to the development and the use of relational databases in general and to the context, structure and uses of the ABCD in particular. There are four sections which could be read individually, depending on the reader's interests or requirements: Sections 1 and 3 are more specific to archaeobotany and Sections 2 and 4 to the practicalities of setting up and using a computer relational database.

1 Background

1.1 Introduction

The intention of this paper is threefold: first, to advertise the existence of the Archaeobotanical Computer Database (ABCD); second to provide some details about the ABCD, its aims, structure, and a brief guide to its use; and third, to provide an introduction to relational databases in general terms and the Structured Query Language (SQL) in particular, for anyone who might be about to initiate a similar database project.

Section 1 explains the reasoning behind the establishment of this project and introduces some of the problems which have had to be addressed during the development of the database. Section 2 is a general discussion of database design and shows the first steps in setting one up. Section 3 gives details of the development of the structure of the ABCD. Section 4 provides an introductory guide to the use of SQL using examples from the ABCD. This is not intended to be a complete 'user guide', however, as one will be given elsewhere (Tomlinson, forthcoming). A glossary of the specialist terms and abbreviations used in this report is presented in an appendix.

1.2 The aims of the project

Research on plant materials from archaeological excavations in the British Isles has produced data, the quantity of which has increased enormously in the last two or three decades. Taken as a whole, these data provide evidence both for the changing distribution and ecological associations of the plants themselves and for the developing exploitation of plants by human communities. Most of this information is either contained in specialist contributions forming chapters in, or appendices to, excavation reports or is held in unpublished archive or laboratory reports where the plant data are often divorced from the archaeological details. Much of it is not readily accessible to archaeobotanists and, not unreasonably, interpretation rarely goes beyond the individual site or regional group of sites (with some exceptions such as Moffet et al. 1991; Greig 1991). There is now far too much information to assimilate and analyse without the use of a computer. One way forward is the creation of a database to hold archaeobotanical records from the whole country so that they can be managed and used in a more effective way (Greig 1991).

The overall aims of the project described in this paper have been to design and develop a comprehensive computer database to store all...
this information in a form which would be useful for several different levels of analysis and could be easily maintained and updated. The requirements of this project were primarily that the database would be used for detailed analysis for the writer's specific research project and, eventually, by any other workers, in a form that would be useful and readily accessible to them.

1.3 'Requirements analysis'

The first stage in setting up a database is to study the data set and to ask the questions 'What information will be required?' and 'In what form will this information be needed?'. This will affect which categories of data are included and in what form they are stored. There are nowadays no real limits on the size of a database which means that it is possible to concentrate on the more important aspects—effectiveness and efficiency.

For the ABCD there were two broad areas of requirements, the first determined by the analyses which archaeobotanists might want to carry out in the future, and the second by the detailed analyses for the author's own research project. Moreover, as mentioned above, the information should be held in such a way that it should be comprehensive, flexible and readily accessible.

1.3.1 Information required by other users

In 1989 the author carried out a survey, in the form of a questionnaire (for further details see Tomlinson forthcoming). The intention was to find out what questions British and Irish archaeobotanists might want to ask of a comprehensive database of the sort that was being proposed and whether they would like to comment on its suggested contents. Issues raised included:

(a) Most archaeobotanists required specific information and asked such questions as:

- What other records (and of what date) are there of a particular taxon?
- What other archaeobotanical sites of similar type and date are there within a given distance of a site?
- What other examples are there of a particular type of assemblage?

(b) Other points made by those who responded to the questionnaire included:

- the quantification or ratios of the abundance of each species should be included.
- the different parts of the plant, for example each cereal chaff element, should be quantified separately.
- some form of quantification of vegetative materials, such as buds and leaves, should be allowed.
- an indication should be given as to the quality of individual identifications of critical taxa.
- information on the archaeobotanist(s) who made the identifications and of the reference collections they used for the identifications should be included.
- details of the location of the data archive and the fossil plant remains themselves should be added.
- for the purposes of association analysis 'sample' species lists would be more useful than 'site' species lists.
- the volume of material processed which produced the recorded remains should be recorded.
- the database should be structured so that it will be easy to select information using several different criteria at once.

Some of these points required various additional categories of information to be gathered; as a result the database is more complex but also more useful.
(c) Certain categories of information which were mentioned by the questionnaire respondents have not been included in the database—partly because of the large amount of extra time involved in searching for the information. Their suggested questions were:

- who was the excavator/site director?
- are the samples narrowly contemporaneous (rather than being only of broadly similar date)?
- what is the spatial relationship of samples on site?
- what are the soil types on the site?
- how does the archaeologist/archaeobotanist feel about the validity of the dating from his/her site?
- could a key-worded bibliography on related subjects such as historical floras and pharmacological information be included?

1.3.2 Requirements of the research project

There are two levels of data analysis which the database is intended to facilitate for this project. First, in broad terms, it is intended to look at the range and occurrence of individual taxa, particularly plants of disturbed ground, to see what types of samples they occur in and if there are any patterns associated with preservation, context type, human activity and so on. This can only be done with a properly relational database containing all the available data for the region and most of the categories of information required by the other archaeobotanists (see (a), (b) in Section 1.3.1), are required for this. Another important requirement is to include Ordnance Survey National Grid references in a form which could be output to a GIS (geographic information system) to produce distribution maps, perhaps with overlays of geographical information such as that relating to soils or geology.

The second level of analysis is to examine certain groups of samples in more detail using relatively simple statistical techniques. The mode of formation and incorporation of biogical remains in archaeological deposits, particularly urban ones, are so complex it might be dangerous to use sophisticated multivariate techniques until some of the broader patterns are examined. To be used for detailed analysis, samples will have to be chosen which were closely dated, had good context integrity (see below, Section 3.1.1, did not consist of materials from very mixed origins and have sufficiently large numbers of taxa to produce reliable results (cf. Kenward 1992, 82).

Various mechanisms are included in the database to allow such 'good' data, of similar quality and compatibility in terms of site type and dating, to be selected. It was thought important to include full details of the methodology used for the extraction and publication of the plant remains—sample weight/ volume, mesh sizes used, quantification, and so on. It was also necessary to develop coding systems for certain categories such as dating and context types, to ensure efficient sorting and retrieval of the right categories of data.

1.4 Theoretical problems

Bearing in mind the two levels of analysis which are required, as outlined Section 1.3.2, there are several problems which need to be taken into account in the design of the structure of the database. These are mainly concerned with the variable and complex nature of published archaeological data.

1.4.1 Quantification

One aspect which makes the comparison of the results given in different published archaeological reports difficult is the quantification of the plant remains. Not only do the sizes of samples and the methods of recording them (i.e. volume or weight) vary but quantification methods also differ. All the plant remains in a sample may have been counted and the numbers quoted in the publication or only the percentage figures may be given. Some 'super-abundant' elements (for example *Laurus* spp., *seeds*) may only be estimated—by counting a subsample and multiplying accordingly. Samples may have been recorded using an abundance scale such as a three- or four-point scale (Hall et al. 1990, 299), or descriptive terms such as 'occasional', 'common', 'abundant' may have been used. Sometimes only the presence of each taxon is noted per sample, or each species is quantified as 'presence, or percentage presence in 'n' samples' (sometimes called 'ubiquity'), see also
Pearson 1989). Often, vegetative fragments are not quantified—only their presence is indicated, even though counts for 'seeds' are given. Because quantification is an important factor in the interpretation of a sample, it is important that this information is recorded in the database using as much detail as is available in the archaeobotanical publication.

In order to carry out any comparisons of samples from different sites, lists of taxa which are required for more detailed statistical analysis can have their quantification converted to a simple three-point scale of abundance. This can take into account not only the proportions of the different taxa in the sample but also the density of seeds in the deposit. The published quantifications would, however, still be stored in the database.

1.4.2 Site dating and classification

In order to look at the temporal distribution of plant species it is important to include dating and site classification information in the database in as much detail as possible, but in a form where different types and dates of site can be easily retrieved.

Although computers are excellent for sorting and retrieving data they are not able to deal with information which is organised in an illogical way, unless this illogicity is in some way accounted for in the design. Because of the nature of archaeological information and the complex systems the human brain uses to 'classify' an archaeological site, it is almost impossible to put this information into a computer, to enable efficient retrieval and sorting, without over-simplifying it.

The data held in the ABCD come from reports published during the last c.100 years. Developments in archaeological dating, both in terms of absolute dating by radiocarbon methods and better understanding of the relationships of post cultures, means that care has to be taken in deciding the method used for recording the date of a site in the database. This is also true for site classification, as terminologies for 'site type' and cultural phases have changed. It is only relatively recently, prompted by the computerisation of Sites and Monuments Records databases, that archaeological thesauri have been developed (for example Leech et al. 1992, Hart et al. 1989). There are several different ways to describe the date or period of a site or context, depending on the nature of the site. Thus, provision has to be made in the database for recording 'date' in several different forms, such as, absolute date, radiocarbon date (either in radiocarbon years BP or 'corrected' absolute date BC/A.D.), dendrochronological date, or cultural period. Although the cultural period is often based on an artefact type, such as pottery, this may not relate directly to any one absolute time or the dating may be different depending on the part of the country i.e. diachronic.

Depending on the accuracy of information from any context the precision of the dating varies considerably within and between sites. Thus one context may be 'prehistoric' but another 'Late Bronze Age' and a third may be described as 'end of the second millennium BC'; yet another may have a dendrochronological date of great accuracy.

There is no simple way of resolving these problems except by recording the information in several different ways in the database, first in as much detail as is available and, second, in a coded form that allows for the selection of sites or contexts using combinations of archaeological period or absolute date. Note that a similar, multiple-field approach for recording dates, has been developed by Smith (1987, 1992).

1.4.3 Sample definition

Although most archaeobotanical reports deal with 'samples' there is no single, clear definition of what an archaeobotanical 'sample' is. A dictionary definition of sample (OED) is 'a relatively small quantity of material ... from which the quality of the mass (or whole), which it represents, may be inferred.' There is rarely any discussion of what a particular 'sample' actually represents and the word tends to be used in an imprecise way. In terms of the mechanics of incorporation of material into a deposit, a 'sample' which consists of a few seeds dispersed thinly through a relatively large layer of inorganic sediment is very different from a 'sample' taken from a discrete accumulation of organic material.

The complex mechanisms by which plant materials become incorporated into archaeological deposits should, of course, never be ignored when the botanical results are interpreted (Kemward 1992, 82).
1.4 Availability of information

With the advancement of archaeobotany over the past three decades, the level of detail quoted in published reports about the archaeobotanical remains has improved greatly. Even so, there are several categories of information which are often not included, for example the provenance of plant material and information about the materials used. Even if it is clear which provenance collection has been used for the identification of a particular taxon, it would be useful to know the range of different taxa actually represented in that collection at the time the determination was made.

Some reports do mention where the archaeobotanical archive is stored but, as this author has occasionally found, this archive may never have been deposited where it is said to be. Information about the sources of material is in which they worked and the institution(s) of which they are members. This information is useful to specialists other than the author of the report in determining the accuracy of the data.

1.5 Summary

The aims and requirements of the project determined the categories of information which needed to go into the database. The complexity of some aspects of the data such as the classification of sites and contexts required careful consideration before the structure of the database was created. The project has highlighted the need for publication of certain details of methodology often omitted from site reports.

2 Relational databases

2.1 Why use a relational database?

Relational databases are designed to cope with large quantities of complex information, to store this information in a form that can be easily accessed, to ensure that it is accurate and that it maintains its accuracy. These points, plus some further ones which are not so relevant to this archaeobotanical database, are discussed clearly by Smith (1992, 114).

Because a relational database is a powerful tool it is most important that the information is structured correctly so that the system can be used to best effect. This is why the database design needs to be discussed in detail (see Section 3 below).

1.2 What is a relational database?

The relational database model was first defined in 1970 when E. F. Codd introduced the idea of using mathematics to represent the structure of database tables and the relations found in set theory as means to model data (Codd 1970). A relational system is one which is constructed in accordance with the group of principles that together form Codd's relational model. One of the main advantages of a relational database over other databases is that the data are stored 'independently' of the database application and not in the much more limiting 'sequential' form. The main principles which affect the construction and manipulation of a database are discussed here (for more detail see Howe 1983; Date 1983; Richards and Ryan 1985; Korth and Siberschatz 1991; Smith 1992).

In essence, a relational database, as viewed by the user, is a collection of tables consisting of columns, for each attribute (e.g. site name) and rows, for each record (e.g. the details of a particular site). The crossing between such column and row is atomic (composed of only one element or cell); see Figs. 1 and 2.

Tables contain information about entities which are things that have properties (e.g. archaeological sites). Entities are linked to each other by entity relationships. Figure 3 shows an entity relationship diagram of part of the ABCD (the boxes represent the tables and the lines linking them the relationships). Thus, an example of a relationship between two tables might be: a table holding information about the entity 'archaeological sites', related to another table with information about a record—archaeobotanical reports on those sites. There must therefore be a mechanism for linking the information held in those two tables. The link is provided by a key attribute which is inserted into both the tables. This key attribute, also known as an identifier might be, as in this example, the site name or a code number representing it.

There are two types of key attribute which link tables. These are called primary and secondary keys (also called foreign keys). Each table has to have a primary key, that is a field,
attribute occurrence or value
sometimes called a 'cell'

Rules that a table should satisfy:
- each column contains values for only one attribute
- the sequence of attributes ACROSS the table is identical; the order of the rows is NOT significant
- every row must be unique, i.e. two rows from the same table MUST differ in AT LEAST one cell.
- all rows must be complete, i.e. each cell should contain a value and that value must be part of the DOMAIN for that attribute (if a DOMAIN has been specified); the value can, of course be NULL.
- repeating groups are not allowed, i.e. only one value is permitted for any cell. (If more are needed it will be necessary to define another attribute, or even another table.)

Figure 1. Diagram of a database table and list of the rules a table should satisfy.

<table>
<thead>
<tr>
<th>site</th>
<th>country</th>
<th>country</th>
<th>grid ref</th>
<th>east</th>
<th>north</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cowick 76</td>
<td>E</td>
<td>HUM</td>
<td>SE</td>
<td>4660</td>
<td>4220</td>
</tr>
<tr>
<td>Danebury 78</td>
<td>E</td>
<td>HMP</td>
<td>SU</td>
<td>4325</td>
<td>1375</td>
</tr>
<tr>
<td>Highgate/Everley 77</td>
<td>E</td>
<td>HUM</td>
<td>TA</td>
<td>5035</td>
<td>4400</td>
</tr>
<tr>
<td>Thorpe Thewles 80-2</td>
<td>E</td>
<td>CLV</td>
<td>NZ</td>
<td>4400</td>
<td>5230</td>
</tr>
</tbody>
</table>

Figure 2. A subset of the sites table as an example of table structure.
or combination of fields which serves as a unique identifier for each of the records in that table. For example, the primary key in the sites table is the site name (site) which links (relates to) the secondary key in the reports table. Primary keys must not contain either duplicated or null (empty) values. Thus any site name must occur only once in the sites table but can be repeated in the reports table.

Information has to be held in these separate sites and reports tables in order to conform with the rules of normalisation (see below).

A relational database language, such as Structured Query Language (SQL), is used to create a database, add data to it and then allow the retrieval and manipulation of the data. This is done by carrying out such operations as: selecting one or more rows from individual tables; joining the information from two or more tables together; or producing a unique list from one field by eliminating all the duplicate rows (for example, if a list of each unique species name occurring in the database was required). Section 4 describes some of these SQL operations in more detail.

The result of any retrieval operation (query) on tables in a relational database is always given in the form of another table (i.e., with rows and columns). The retrieval operations—such as the SELECT statements described in more detail in Section 4—have to conform to a set of operators, derived from relational algebra, which form the manipulative part of the model. Sets or subsets of information are retrieved from the rows and columns of one or more tables of the database and the results constructed to form a new table (see for example Boxes O1—O5 in Section 4).

There is a confusing array of terminology in database management theory both for the elements of databases and the development of them. There are often two or three words for the same things (e.g., row/table/record) and it seems that each new book that is published uses its own terminology. This is presumably because this is a relatively new subject. A word is chosen to represent something which, unfortunately, is not entirely appropriate. For example 'atomic' which has been used to describe the contents of a database table cell (see above) is confusing because an attribute CAN be composed of several words. It is difficult to tell at this stage which wording will become the conventional one. In the previous paragraphs the words 'column' and 'field' have both been used and, although the words do refer to the same thing, there is a subtle difference in their use. A glossary of terms mentioned in this paper is given in the appendix.

2.3 Normalisation

Another feature of a relational database is that it should not have any unnecessary sets of duplicated data. This is important—particularly in complex multi-user systems which are continually being updated. It may not seem so important in a rather static database such as the ABCD but mistakes are less likely to be made if the structure conforms as closely as possible to the true relational model. Removal of redundant sets of data and streamlining the structure of the database comes under the process called normalisation which has developed out of the theoretical approach to databases (Date 1986).

Normalisation is the process of looking at the different types of relations between, and within, the entities in a proposed database. The relationship between the entities 'archaeobotanist' and 'archaeobotanical report', for example, is described as many-to-many (m:n) because an archaeobotanist may have written several reports and any one report can...
Figure 4. Entity relationship diagram of part of the ABCD: (a) before normalisation; note the m:n relationship between reports and archaeobotanists (workers); (b) Final structure. Primary keys (see Section 2.2 for explanation) are underlined.
have several authors. The relationship between the entities 'sites' and 'contexts' is one-to-
many (one site can have many contexts, or a context can originate from one site, 1:n), but the
between 'a site' and its unique 'grid reference' is one-to-one (1:1). To conform to the rules of
normalisation, a database should not have tables with many-many relationships. Thus,
for example, two tables which are found to have a many-many relationship have to be
'normalised'. Another table is inserted to link them, thus creating two 1:n relationships (see
Fig. 4). This is why the worked table in the ARCD links the reports (archaeobotanical
reports) and who-who (archaeobotanists) tables. Thus each item of information held in
the worked table is only listed once and, most importantly, only has to be updated in one
place—for example if further information about a worker needs to be added). The worked
table represents the entity 'archaeobotanists who worked on reports'.

A further stage of the normalisation process also removes redundant data. Attributes may
be found to be functionally dependent on one another within an entity relationship. For
example, say a 'sites' table had a field 'county' and a field 'county_code'. In this case, 'county'
would be fully dependent on 'county_code'. There is no point repeating the county name
each time in the 'sites' table; a separate table with a unique list of the full county names
and their codes would be more efficient. Only the primary key should have attributes which
are functionally dependent on it. Attributes that are dependent on other non-key attributes
are best removed to another table unless there is a good reason to keep them. For a
further example of a functionally dependent attribute see Section 3.2. To summarise: a
relational table is said to be fully normalised if every attribute that is not part of its primary
key is dependent upon that key and is not dependent on anything else in the relation.
(For further straightforward explanation see Howe 1983.)

There are instances where the rules of normalisation may be broken, particularly if it
is more efficient and convenient to do so without causing problems. An example of this is if
one attribute of a 1:m relationship only repeats very occasionally. Thus the relationship
between reports and bib (as shown in Fig. 4) is 1:1. Each record in the bib table relates to
one record in the reports table. Occasionally, however, a report contains information on
more than one archaeological site (e.g. Creig

1988) which means that some of the
bibliographic information has to be repeated
which, strictly, should not be so.

2.4 The conceptual data model and
database design

The conceptual data model is defined as 'some
abstract and general description of that part of
the universe which is to be represented by
data in the database' (Papazoglou et al. 1989;
for further explanation relevant to archaeology
see Burnard 1992).

After the 'requirements analysis' is completed
(see Section 1.3 above) and decisions have
been made about the data required and the
form of output needed, the next stage is to
produce a 'conceptual data model', as follows:

2.4.1 Decide what the entities are: These may
be, for example, 'plant identifications',
'archaeological reports', 'bibliographic
references' 'archaeological sites'. Then decide
what the relationships are between these
entities, for example, 'plants found from sites',
'reports written about archaeological sites' and
so on.

2.4.2 Name the entities: Names (identifiers)
need to be chosen for the entities. Should
'sites', for example, should be identified by a
reference number (which would be a simple
method) or by their actual names? In the
present project having started by using
several numbering systems to identify things
it was found to be easier to cope with actual
names, especially while collecting the data. It
is also useful at this stage to think carefully
what an entity actually represents and choose
a name for the table which is unambiguous
but easy to use and remember.

2.4.3 Decide what relationships there are:
Decide what relationships there are and of
what type. Think carefully about what
categories of information will be required
from the database as this may help to decide
where, and in what form, information is
stored.

2.4.4 Expand many-to-many relationships:
Make sure there are no n:m relationships by
creating extra tables. The insertion of the
worked table (see Section 2.2) is an example of
this.
2.4.6 Remove functional dependencies: Remove relationships within tables where fields are functionally dependent on other non-key fields instead of redundant data (see Section 2.2).

2.4.6 Decide on the attributes for each field: Look at the attributes and define sets of values for 'attribute occurrences'. Decide whether to create 'domains' of the attributes for any one entity. A domain is a set of values for an attribute; each attribute occurrence must be one of the values of the domain. In reality this often produces simplification which is not necessarily a good thing, it is sometimes better not to set parameters that are too rigid, especially while developing the database and collecting the data. (For further discussion on this see Section 3.1.1 below).

2.4.7 Look for and remove sub-types: Consider whether there are any sub-types. If these are really two entities combined in one field it may be better to split these into two fields. For example, a general site description category in the sites table had to be split to take account of the condition of the site at the time of excavation as well as the topography of the site at the time of deposition. The advantage of the relational system is that fields can easily be added later, but obviously it is important to avoid having to go back and look through all the references again because insufficient detail was collected from the start.

2.4.8 Create the database: Once the entities and relations have been defined, the database can be created—the entities become the tables and the attributes become the columns. The relationships between the tables are then established by creating primary and secondary keys. Other column constraints such as specifying UNIQUE and NOT NULL fields (those which must not contain null values) or particular data types may also be defined. Inverse will also need to be created to speed up retrieval operations.

2.4 Practical questions about hardware and software

It is difficult to choose which database management systems are the best to use because of the very rapid development of computer software and hardware. There seems little doubt that a relational database system which uses SQL is sensible because it is an international standard and systems which use it are available on both mainframes and microcomputers. Note that SQL is a database language not a management system, but it is the language used in many relational database management systems (DBMSs).

The ABCD has been set up on the York VAX mainframe computer, which has SQL, but no proper data management system. This is adequate for the purposes of this research project but in order to produce an easy-to-use 'front-end' for other users and for improving data checking and integrity systems, it will be necessary to have the database in a DBMS such as ORACLE.

In the questionnaire survey in 1989, archaeobotanists were asked which computers and packages they used for their work. As to single database package emerging as the most often used, it seemed that it would not be feasible to distribute the database as a complete package (see also Booth et al. 1989). The ABCD is to be maintained and used centrally at the Environmental Archaeology Unit where queries and updating can be carried out and an annual catalogue produced if sufficient subscriptions for this can be found. The timing of the up-grading of the York University VAX mainframe to a network system has also added complications to this, as it is not clear whether it will be feasible to access the database via the network. Creating an on-line package accessible to remote users would require more effort than it is probably worth expending, especially as only a few archaeobotanists have access to the academic communication network (IANET).

2.5 Summary

The effort required to understand the somewhat complicated background of relational databases may seem unnecessary but it is hoped that in doing so a well-structured and truly relational database will be the outcome. If this is the case, then the information held in the database is more likely to be accurate, less likely to become corrupted and much more easy to retrieve in a complex form.

3 ABCD Database design

The development of the archaeological computer database (ABCD) has had various
constraints. First, the information collected needed to be of sufficient detail and complexity for it to be useful in answering a fairly wide range of questions. Second, the material had to be stored in such a way that it would be easy to access and to transfer to other database systems if necessary. Third, inputting of data had to be flexible so that data could be collected on various computers and transferred to the database.

3.1 Development of the ABCD

Only categories of information which require some explanation or for which difficulties have been encountered during data collection will be discussed here. The full list of tables and their fields is given in Section 3.3.1.

One of the main difficulties with such a large and variable data set is to construct the database so that it is possible to select useful and meaningful subsets of the data. There is a need to include as much detail as possible but at the same time have some coding systems which allow simple 'select' statements to be used (see Section 1.4.2 for further discussion of this). First, there is the problem of classifying things, such as type of archaeological site, in a way that will make sorting straightforward. Second, there is a need to be able to select 'good' data. There should be some means of providing adequate 'quality control' over the information in the database.

3.1.1 Site classification—The ABCD has a column of variable length field for site classification (type). The terminology used in the archaeological report (in as many words as necessary) is recorded here. Over most of the data have been entered into the database a certain amount of rationalisation of the wording used can be carried out. A thesaurus can then be produced with a cross-referencing to alternative words or synonyms. This follows the 'bottom up' approach as described by Chadburn (1988). There also needs to be a distinction between the classification of the

<table>
<thead>
<tr>
<th>1st Level:</th>
<th>2nd Level:</th>
<th>3rd Level:</th>
</tr>
</thead>
<tbody>
<tr>
<td>PH Prehistoric</td>
<td>ML Mesolithic</td>
<td>BE Basket (Neolithic or early IA)</td>
</tr>
<tr>
<td></td>
<td>NB Neolithic</td>
<td>EB Early Bronze Age</td>
</tr>
<tr>
<td></td>
<td>IA Bronze Age</td>
<td>MB mid Bronze Age</td>
</tr>
<tr>
<td></td>
<td>1A Iron Age</td>
<td>LB late Bronze Age</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SB Bronze Age/Iron Age transition</td>
</tr>
<tr>
<td>RO Roman</td>
<td>RM Roman military</td>
<td>EI early Iron Age</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MI mid Iron Age</td>
</tr>
<tr>
<td></td>
<td></td>
<td>LI late Iron Age</td>
</tr>
<tr>
<td>PC Pictish/Celtic/early Christian</td>
<td>BB Roman non-military, civilian</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>ER early Roman c.3rd-2nd century AD</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MB mid Roman c.2nd-1st century AD</td>
</tr>
<tr>
<td></td>
<td></td>
<td>LX late Roman c.3rd-4th century AD</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PR post-Roman</td>
<td>SE Saxons/early medieval</td>
<td>ES early Saxons/Dark Age AD 400-650</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MS middle Saxons incl. Anglian, Pictish AD 650-850</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AS Anglo-Saxons/Scandinavians Saxon-Norman AD 850-1150</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MB early medieval AD 1150-1250</td>
</tr>
<tr>
<td></td>
<td></td>
<td>HM high medieval AD 1250-1400</td>
</tr>
<tr>
<td></td>
<td></td>
<td>LM late medieval AD 1400-1500</td>
</tr>
<tr>
<td></td>
<td></td>
<td>16 16th century</td>
</tr>
<tr>
<td></td>
<td></td>
<td>17 17th century</td>
</tr>
<tr>
<td></td>
<td></td>
<td>18 early 18th century</td>
</tr>
<tr>
<td>MO modern, post AD 1750</td>
<td>ND not dated</td>
<td></td>
</tr>
</tbody>
</table>

Figure 5. Date coding system.
site at the time when the material was deposited and at the time when it was excavated. Thus the field typing is for recording the typography or geographic location at deposition and unit for recording the unit of the site at excavation (e.g. under water).

3.1.2 Workers: The worked and whoswho tables allow both the person(s) who did the archaeological work and any person(s) who helped or confirmed the identifications of taxa to be recorded even if they are not the author(s) of the published report. Alternatively, for determinations of individual taxa, a note can be added to the probs table.

3.1.3 Sampling: The samples table has fields for recording the methodologies used at the different stages: taking the sample from the ground; processing to extract the plant material; recording the taxa in the publication. Lists of taxa from groups of contexts of the same date range and type, may be combined in the database if there is no good reason for keeping them separate.

3.1.4 Context integrity: This is a means of selecting 'good' data using a three-point scale (see Section 3.2). The score is high (3) if a sample is from a well-sealed context, representing material deposited at one time, or over a relatively short period, implying that the risk of contamination will be low. The integrity score is low (1) if the archaeological information is poor or if it is clear that the context contains material derived from several taphonomic processes. It should be understood that the scoring will be fairly subjective as it is based on an interpretation of the evidence in the archaeological report by the database compiler.

3.1.5 Dating security: Similar to context integrity, this three-point scale indicates how well the context is dated. Again, this is subjective and it is based only on the available published information and the opinion of the database compiler. The score given will depend on the dating methods used and the likelihood of the date range that is quoted being 'acceptably' accurate. The dating table allows details of dating methods and radiocarbon dates to be recorded, where the information is available.

3.1.6 Dating codes: To enable information to be selected from the ABCD either in broad terms, such as 'Roman' or 'post-Roman' or where the dating is more precise, as for example 'late Bronze Age' or 'Anglo-Scandinavian', there is a table of the system, held in a separate codes table (see Fig. 5). The information held in the codes table will be two entries for each type, particularly of contexts dated more precisely by radiocarbon methods, but will only be used for grouping together the context_lists of similar date. More details about the dating is also held in the contexts table (see Section 3.3).

3.1.7 Taxonomic problems: There is provision in the database (probs table) for storing information about each individual plant identification. This covers several categories. It serves both as a quality check and as means of surmounting such difficulties as those created by the updating of taxonomic names. Where there are variations in the identification names which archaobotanists use for a group of taxa (for example: Ranunculus Section Ranunculus; Ranunculus acris/lubusus/repens; Ranunculus repens type) these can be recorded separately.

The probs table holds the information about individual identifications. This means there need be no restriction about the additional information stored about an entry in the taxa_lists table. The information is linked to the taxa_in the taxa_lists table by a code. By using this code it is possible to see from the taxa_lists table what type of problem it is. If it is only minor, such as a change in the taxonomic name, then the probs table itself need not be referred to. The fields which relate the taxa_lists and probs tables are list and prob.

The problem codes used so far are as follows:

A where alternative names have been given in the report these can be listed.
B if the material is possibly a contaminant (original author's suggestion).
C if the identification is doubtful—in the opinion of the database compiler.
D where the quantification of a taxon is an estimated figure derived from counting only part of a sieved fraction.
E if the naming level was too precise in the light of current knowledge.

12
where different taxonomic names have been used which makes the identification unclear; the name given in the report should be given, with an explanation if necessary.

if species are indicated as spp., or sp., i.e. more than one species is thought to be present although they cannot be identified separately.

if the quantification is not the actual numbers of individuals, but, for example, given as a percentage.

if there are several problems relating to this taxon, listed in the probs table.

each taxon has been used from the ARCD checklist; but there is no problem with the identification.

If the taxon is not on the checklist and therefore needs to be added.

where only common names were used in the report; the most appropriate Latin name would be entered into the taxon_lists table and the English given in the probs table if there is a possibility of any confusion.

Quantification: The quant field in the taxon_lists table takes numbers up to 9999 for a quantification symbol such as 'a' for abundant, 'p' for present, or any other system used in the published report. Explanations of these symbols are then given in the probs table. A further field can be added to the taxon_lists table to take a three-point abundance scale for lists which require semi-quantitative statistical analysis.

Individual identifications: Archaeobotanical reports which were not published recently and have no detailed sampling and context information or reports which are simply individual identifications and not 'proper' samples are indicated in the sites table with an asterisk. For these reports no information would be added to the samples and contexts tables.

Other reports: Other reports on biological remains noted in a publication are recorded in the report table. Some workers record plant material in their plant lists which are identified only as 'parts' such as twigs, buds and leaf fragments. These can be recorded in the same field of the taxon_lists table as 'indet.' and an explanation given in the report field. Non-plant material such as eggshell or insect fragments can be recorded in a similar way. These elements cannot be given in the name field as only plant names which are listed in the checklist (or 'indet.' should be in this field.

Missing data: Where information is not available from the published report or is not relevant 'or' should be used as it is better not to have empty attribute occurrences in any field. Where there are missing data which have yet to be filled in, for example where the grid reference is missing, question marks should be used. Several of the fields have been set up with column constraints 'NOT NULL' which ensures there are no empty cells.

Interpreted information. The intention has been to record information as it is presented in the published reports. However, where the information recorded is an interpretation by the database compiler this needs to be clearly indicated.

Bibligraphic references—bibliographies are notoriously difficult to put into a database because to fit into a good relational system each reference has to be divided into a number of different tables. Allowance has to be made for the various levels of nesting, for example where a botanical report is in an archaeological report which is in a chapter in an edited book, or part of a series. With such a database system it would be possible to search and sort the bibliography by author, book title, publisher, date, report title and so on. This is not, however, needed for this project, especially as it would have added a considerable amount of time and effort to set it up. By using the WordPerfect word-processing package and a text file with each reference stored as a separate paragraph, it has been found to be easy to handle, search and sort the bibliography. All or part of the bibliography can be printed out at a moment's notice, in a properly formatted form. The disadvantage of this system is that there is some duplication of references but it was decided that simplicity was more important than space saving in this case.

By following a few simple rules of lay-out, the bibliography can be sorted into order by author's name or journal name to facilitate
efficient searching in the library. The report number is added to each reference when the
information from that report is recorded. Thus a finished bibliographic record has the format
shown in Figure 6. Each reference is in a
paragraph separated by a blank line.
With this format WordPerfect’s ‘paragraph sort’
can be used to order the references by, for
example, the first word (‘word number 1’) of
the second line of the paragraph—to give
author order, or by the last word (‘word
number —1’) of the second line to give
publication date order. Obviously, the latter
example will only work if the second line
always ends with the publication date. In
WordPerfect up to nine of these word
categories can be sorted in one operation.

During data collection two bibliography text
files, ‘bib.bib’ and ‘fin.bib’, have been used.
Unrecorded references are entered into bib.bib
and they are transferred to fin.bib (using
WordPerfect’s BLOCK, COPY and SWITCH
commands) and given a report number when
the information is entered into the database.

Full stops have been used to separate elements
of the bibliography so that at a later stage
‘fin.bib’ could be loaded into a database
requiring a more complex structure, if anyone
wished to take the time to do so.

3.3.1 Data collection: To facilitate maximum
flexibility when collecting and inputting the
data, files have been word-processed (using a
portable machine) and later ‘input’ into the
database using a loading program.

3.2 Comments on the ABCD structure

Although the ABCD is essentially fully
normalised (cf Section 2.1.1) there are some
exceptions. For example, in the sites table
gridref (the two-letter 100 km square reference)
is fully dependent on two other fields, east
and north.

Normally a field dependent on another non-
key field should not be in the same table (see
Section 2.3). It would be possible, rather than
storing the grid square codes in the sites table,
to generate them only when they are required,
by using another table containing a unique list
of the codes. Gridref has been retained in the
sites table because it seemed simpler and was
a way of double checking that the references
were correct. The grid reference is required in
co-ordinate form for loading into a
distribution map program or Geographic
Information System (see Section 4.3) but
people are generally more familiar with the
letter-code system (see Fig. 7).

A six-figure grid reference in the form:
(where £ = numerical grid square reference 43)
given as a co-ordinate needs to be in the form (easting,northing):

| SK 255655 | 4255,5655 |

Figure 7. Grid co-ordinates.
Figure 8. Diagram showing the relationship of the ABCD tables. Key: Boxes represent database tables; table name (bold); primary key (underlined); lines represent relationships; dashed lines show additional/possible relationships (see Fig. 4 for a key to the lines linking the boxes).
The information about sites and reports is not held in one table (which might be thought to there may be more than one report for any one site for example a 'charcoal' report and a "macrofossil" report).

3.3 Structure of the ABCD and details of the tables, attributes and relations

Although the primary entities of the ABCD are archaeological sites, reports, samples and plant identifications, after the processes of database design (outlined in Section 2 above) the final structure of the database consists of at least 19 tables. These are shown in Fig. 8. Each box represents a table, the primary and foreign keys are shown and the lines represent the relationships between the tables. Further tables are likely to be added to the database in due course, such as various coding systems which will refer to the plant catalogue and be used to analyse, for example, the ecological groups represented in the samples.

Figure 9 shows the recording sheet which was developed for collecting the data from each archaeobotanical report. It has continuation sheets for recording further contexts and samples for each taxon list to be recorded in the database (not shown).

The following section (3.3.1) gives a full list of the columns in each table with some notes on their attributes, further details of which will be provided elsewhere (Tomlinson forthcoming).

3.3.1 List of ABCD tables and their attributes

The tables are listed in alphabetical order followed by their column names with a brief description of the attributes (some of these have already been discussed more fully in Section 3.3).

<table>
<thead>
<tr>
<th>Column</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>arch</td>
<td>Archive for archaeobotanical material.</td>
</tr>
<tr>
<td>arches</td>
<td>Provides a list of all archives of archaeobotanical material referred to in the database.</td>
</tr>
<tr>
<td>archr</td>
<td>Archive code number.</td>
</tr>
<tr>
<td>archesr</td>
<td>Details of the archives.</td>
</tr>
<tr>
<td>bibs</td>
<td>Gives the bibliographic reference for the report.</td>
</tr>
<tr>
<td>report</td>
<td>Report number.</td>
</tr>
<tr>
<td>reference</td>
<td>Full reference, in standard Harvard format. Abbreviations of journal or series names follow the British Standard (BS 444).</td>
</tr>
<tr>
<td>checklist</td>
<td>Gives a full, unique list of plant names which occur in the database, with codes to enable sorting or cross-referencing with other databases.</td>
</tr>
<tr>
<td>name</td>
<td>Taxon.</td>
</tr>
<tr>
<td>family</td>
<td>Family.</td>
</tr>
<tr>
<td>genus</td>
<td>Genus.</td>
</tr>
<tr>
<td>species</td>
<td>Species.</td>
</tr>
<tr>
<td>sort</td>
<td>Sort code number to enable inter-species or grouped taxa to be printed out in taxonomic order.</td>
</tr>
<tr>
<td>certain</td>
<td>Certainity code number distinguishes identification levels such as 'sp.' and 's.f.'.</td>
</tr>
<tr>
<td>codes</td>
<td>Provides a means of setting on context type and date using abbreviated codes (the context table provides more detailed information).</td>
</tr>
<tr>
<td>taxon</td>
<td>Taxon list number (see TAXONLIST table).</td>
</tr>
<tr>
<td>type</td>
<td>Code for context type.</td>
</tr>
<tr>
<td>per</td>
<td>Code for cultural period (see Fig. 3).</td>
</tr>
<tr>
<td>cell</td>
<td>Code for reference collection.</td>
</tr>
<tr>
<td>info</td>
<td>A complete listing of the full name and location of reference collections used in the identification of the plant remains.</td>
</tr>
<tr>
<td>colls</td>
<td>Lists REPORTS table to the details of the reference collections used in the COILS table.</td>
</tr>
<tr>
<td>report</td>
<td>Botanical report number.</td>
</tr>
<tr>
<td>coll</td>
<td>Code to reference collection used.</td>
</tr>
<tr>
<td>contents</td>
<td>Gives details of the context of the sample(s).</td>
</tr>
<tr>
<td>list</td>
<td>Taxon list number (see TAXONLIST). Context multiplicity as given in the report.</td>
</tr>
<tr>
<td>number</td>
<td>Context type description, in a minimum of number of words, e.g. PP 3 (1).</td>
</tr>
<tr>
<td>integrity</td>
<td>Context integrity code, on a scale of 1 to 3.</td>
</tr>
<tr>
<td>0</td>
<td>Controversial.</td>
</tr>
<tr>
<td>1</td>
<td>Low—the material is not from a well defined context, the archaeology is not particularly clear, or there is doubt about the sampling method; samples may not relate to the archaeology's context, layers or phases.</td>
</tr>
<tr>
<td>2</td>
<td>Medium—the context is from a clearly defined feature or features and the archaeology is clear but there is no...</td>
</tr>
</tbody>
</table>
indications that the context was particularly well sealed. There may be evidence from artefacts, for example, that the context contained re-worked materials.

3 High—It is clear that the context is well sealed and closely defined and the archaeological evidence is very strong.

4 New: Main preservation mode of the plant material: low-resolution mosaic

5 Sedimentary or 'well' description, quoting directly from the report.

6 Cultural period / date for the context, using the terminology given by the excavator:

begin begin: the earliest likely date of the context, given absolute dates, using a minus-Ieg for years BC. Details of the date quoted on the report can be given in a separate dating table. Thus, 450 BC—AD 750 would be input as begin date: 450 end date: 750. Where a radiocarbon date is used, this date range will be given as 'within ± standard error which is not far from the true 95% confidence limits. That '400±20 BC would be quoted as begin date: -400 end date: -120'. This would be using calibrated dates where possible.

7 Method: Main dating method used for the context—the dating evidence (e.g. C14 dates) may not necessarily come from the actual layer which was sampled, but sufficiently near to it to be related or associated.

8 Validity: Validity of the dating evidence (e.g. C14 dates) often being a three point scale. Note that this will often be an interrelationship of the evidence given in the report.

9 Poor accuracy of dating—how the dating has been determined using stratigraphic relationship alone and the limits are only available.

10 Middle—the dating has been determined using more than one independent method (e.g. stratigraphy plus artefact evidence) and the limits are reasonable. If there is a radiocarbon date, but only from associated contexts, the validity would be medium.

11 High—the dating has been determined using good stratigraphic and artefact evidence or good radiocarbon dating, or preferably both. The probability that the plant material in the begin and end date limits is very high. Contexts with rarely be given this score.

12 Environmental interpretation, maintained given the report.

arch Archaeological interpretation notes about the interpretation as the context from the archaeological evidence

county or county site—unique list of county codes and county names in full, to relate to sites table.
county or county name

dating gives details of any radiocarbon dates, relating to any taxon list.

list Taxon list number (see TAXON LIST table).

note: Quote the C14 dates and details. The date may refer to the whole site or the context concerned.

amps provides an indication of the existence of other environmental reports in the same, or other, publications.

parts full list of the stratigraphic divisions used in the taxon site table for each plant.

part part letter code and taxon diagnoses table.

part name Explanation of the part code.

probes provides a means of detecting any problems or additional information about any one identification.

list Taxon list number (see TAXON LIST table).

note Problem code.

report Details about the problem, for example by taxonomic name used in the report, a note about contamination, the level of identification or an explanation of the qualification codes.

reports gives details of each archaeobotanical report.

report A four-digit code for each botanical report. This number is used to relate several of the other tables to each other and the SPITB table.

site Site name—relates to site in SPITB table.

pal Plant: Ynt the wok was published.

type The type of report—what material was analysed for the report. One letter code: c = wood (charred or waterlogged), g = 'spot' identifications of cereals (grains or impressions).
### 3.4 Problems not addressed

There are several aspects which have not been dealt with in this database; for example, there has been no attempt to relate the botanical information to other biological results from the same contexts, apart from noting the presence of other environmental reports (e.g., Table 1). It would have been impracticable to attempt to include negative information, such as sites which did not contain plant material (even though it was looked for) or samples which did not contain any plant materials (although these may have been listed or accounted for in the published report). It should also be borne in mind when carrying out any analysis on this material that there are large gaps in the data and thus some questions are not answerable at this stage.

Contemporaneity and spatial relationship of samples on site are two aspects very rarely covered in archaeological reports and so it
has not been possible to include these categories in the ABCD.

Because of limitations of time, where more than one sample had been recorded from one context or group of similar contexts, of the same type and date, these have been amalgamated. It would, however, be quite possible in such cases to add separate lists of taxa to the database at a later stage if they were required for some analysis.

Time limitations have also prevented the inclusion of: (a) pollen reports (except where pollen was analysed from macrofossil samples); and (b) records from elsewhere, such as other European countries, although there is no reason why these could not be added later.

It was also decided not to include unpublished reports in the database, partly because of the difficulty of obtaining the archaeological information. The majority of Ancient Monuments Laboratory reports fall into this category.

One minor difficulty in the data collecting phase of the project has been making sure that: (a) reports are not recorded twice by mistake; (b) references dealing with the same site are cross-referenced to; and (c) the most detailed report is used rather than a summary or inform report. This requires careful cross-checking which could perhaps be developed into an automatic system in the database in the future (see Section 2.4).

3.5 The elements of a 'good' archaeobotanical report

Many of the hundreds of archaeobotanical reports which have been examined for this project might have been written or structured in a way which would have made them more useful to the reader and some suggestions for standardisation, simplification and improvement are discussed here. The author is fully aware of the difficulties associated with publishing specialist reports—difficulties relating to the time it often takes for them to be published, editorial alterations made without consultation with the specialist, limitations on report length, non-availability of final proofs and sometimes the lack of proper consultation with other specialists. Despite these difficulties, certain guidelines could be followed in the writing of reports which would enhance their usefulness.

One of the most vital points is to the link the botanical results to the archaeological interpretations. At the very least, page references to the specific relevant point in the archaeological report are essential (for example, where each context is described, listed or illustrated). Details of dating, context type, sediment condition and so on should be easily cross-referencable, or given in detail in the botanical report. One problem that arises when reports are not properly integrated is that the archaeological information and bibliography become separated from the botanical report when it is printed. A second difficulty occurs when the methodologies used for the processing, recording and reporting of the plant material are complex. The use of the word 'sample', for example, can be very misleading. If the published list of plant remains contains some actual numbers and some estimated numbers of plants from a sub sample or sieved fraction, or where certain 'samples' lists presented within one table have been obtained by different methods, perhaps the use of a flow-chart might make it possible to standardise the explanation of methodology and to make it more straightforward to comprehend?

Another problem concerns the level of identification of taxa which is sometimes given differently in a table or list of taxa to that which is mentioned in the accompanying text. For example, where a list gives Fraxinus sp., but the text discusses T. spicata and T. discoidea. Clarity in explanation of methods used for quantification of plant remains is important, for example, where 'fractions' of a sample have been treated in different ways. It is also important to quantify vegetative material at least on a scale of abundance (Hall, et al, 1990, 293), so that it can be related, at least in broad terms to the abundance of seeds.

Needless to say, the author hopes that archaeobotanical reports will contain all the relevant information to make it easy to fill in the recording form (Fig. 3) for updating the ABCD. Report writers might consider using this form as a 'check-list' when preparing their reports. It is easy to omit certain useful categories of information such as details of the reference collections used or the final intended location of the archive and the fossil material.
<table>
<thead>
<tr>
<th>Column</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>SITE NAME &amp; EXCAVATION DATE</td>
<td>SHEET NO. OF</td>
</tr>
<tr>
<td>OTHER NAME</td>
<td>ACCESSION CODE.</td>
</tr>
<tr>
<td>ARCHAEOLOGICAL SITE CATG.</td>
<td>COUNTY. COUNTRY.</td>
</tr>
<tr>
<td>EXCAVATION NOTES</td>
<td>GRID REF.</td>
</tr>
<tr>
<td>SITE PERIOD RANGE</td>
<td>TOPOGRAPHY OF SITE.</td>
</tr>
<tr>
<td>BOTANICAL REPORT NO.</td>
<td>DESCRIPTION OF SITE.</td>
</tr>
<tr>
<td>WORKERS' INITIALS</td>
<td>PUB. DATE.</td>
</tr>
<tr>
<td>REF. COILS</td>
<td>REF. COILS</td>
</tr>
<tr>
<td>REPORT FULL REF.</td>
<td>REPORT FULL REF.</td>
</tr>
<tr>
<td>CONTEXT NO.</td>
<td>TAXON LIST NO.</td>
</tr>
<tr>
<td>CONTEXT TYPE</td>
<td>SAMPLE NO.</td>
</tr>
<tr>
<td>CONTEXT INTEGRITY: 0:1:2:3</td>
<td>SAMPLE TYPE: STANDARD BLANKET COMBINED</td>
</tr>
<tr>
<td>PRESERVATION:</td>
<td>SAMPLE SIZE.</td>
</tr>
<tr>
<td>SEDIMENT DEC.</td>
<td>SAMPLING METHOD.</td>
</tr>
<tr>
<td>CULTURAL DATE</td>
<td>DISAGG. METHOD.</td>
</tr>
<tr>
<td>DATE BEGIN END</td>
<td>MESH SIZE.</td>
</tr>
<tr>
<td>DATING METHOD</td>
<td>SORTING METHOD.</td>
</tr>
<tr>
<td>DATING VALIDITY: 0:1:2:3</td>
<td>TOTAL NO. OF TAXA.</td>
</tr>
<tr>
<td>ENV. INTERPRET.</td>
<td></td>
</tr>
<tr>
<td>ARCH. INTERPRET.</td>
<td></td>
</tr>
<tr>
<td>TOTAL NO. OF RECORDED CONTEXT</td>
<td>SAMPLE LIST: TAXA.</td>
</tr>
<tr>
<td>CHECK: AUTHOR: ARCHAEOLOGICAL OTHER REF:</td>
<td>ALL SAMPLES: RECORD COMPLETED.</td>
</tr>
<tr>
<td>SELECTED NO.: SHEET INPUT:</td>
<td></td>
</tr>
</tbody>
</table>

Figure 9. Recording sheet.
query operations, it can be used as a data sub-language within a larger application (such as a program written in FORTRAN, C or Pascal). This is usually called 'embedded SQL'. Details of the SQL FORTRAN package used for this database will be made available elsewhere (Transimion forthcoming). Thirdly, and becoming more common, SQL is used as the command language in database management packages such as ORACLE, ENCORES and dBASE IV. Again, within these packages it may also be used either interactively or embedded in a program.

4.2 Some examples of 'queries' and output

Included here are some examples of SELECT statements and their resultant output, which should provide an indication of the flexibility of SQL and of the ease of manipulation of the data held in the ABDC.

In these examples, the SQL queries (SELECT statements) will be found in Boxes S1, S2 and so on, and examples of their associated output in Boxes O1, O2, etc. The output, for the purpose of this paper, has been transferred to a word-processing system where some codes affecting layout have been added.

### Box S1: Find all the reports in the database published in 1956

**SELECT** site, type, report, pub
**FROM** reports
**WHERE** pub LIKE '56';

### Box O1: Results of the query shown in Box S1

<table>
<thead>
<tr>
<th>SITE</th>
<th>TYPE</th>
<th>REPORT</th>
<th>PUB</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aclandton Camp</td>
<td>c</td>
<td>1238</td>
<td>56</td>
</tr>
<tr>
<td>Pipton Long Cairn</td>
<td>c</td>
<td>0158</td>
<td>56</td>
</tr>
<tr>
<td>Hunt House 5+5</td>
<td>i</td>
<td>1107</td>
<td>56</td>
</tr>
<tr>
<td>Sarlshof 25+35</td>
<td>w</td>
<td>1147</td>
<td>56</td>
</tr>
<tr>
<td>Thrilow 53+4</td>
<td>c</td>
<td>1430</td>
<td>56</td>
</tr>
<tr>
<td>Thrilow 53+4</td>
<td>g</td>
<td>1431</td>
<td>56</td>
</tr>
<tr>
<td>Leigh Freughan 51-2</td>
<td>g</td>
<td>1781</td>
<td>56</td>
</tr>
</tbody>
</table>

7 rows selected

It is possible to put some layout commands in the SELECT statement—see Box S3.

Box S1 shows a SELECT statement in its simplest form. The output (Box O1) shows that there are only six reports from 1956 in the database and two of these are from the site (Thrilow 53+4)—one an identification of charred cereals and the other a wood or charcoal report. Boxes S2 and O2 show a SELECT on the sites table where * indicates that all columns are to be selected. The IN predicate is more useful than LIKE in this instance as a list of items to be selected can be given (in this example the items are report numbers). WordPerfect layout commands have been added to achieve this particular output format.

Boxes S3 and O3 show a query on the sites and reports tables of the database. The line

WHERE site = reports/site relates the two tables. The query asks for an alphabetical list of all site names found within the 100 km grid squares (SP and SK), where the report number is between 1000 and 1060 and exceeding records for which period contains a 'P'.

If there are two fields with the same name in two different tables (as there logically often have to be then in a SELECT statement, in order not to be ambiguous, the table name has to be quoted as well. In Box S5, for example, the two site columns have to be distinguished (e.g. reports/site and sites/site). This can also be achieved by giving the table name an abbreviated alias, as in Box S5 (FROM contexts c, taxon_lists t).

The output shown in Boxes O2 and O3 gives examples of the information that could be provided in the catalogue that is intended to distribute to subscribers to the ABDC.

The CREATE statement can be used to set up views. These are 'virtual' tables, derived from a relation of other tables, holding a subset of data, but the data are not physically stored in them. These are useful when the database gets very large and one is temporarily only interested in certain aspects of it. A view is created in Box S4a to hold the number obtained by COUNT so that this table can then be 'queried' to sort the numbers into descending order (Box S4b). Note the two columns in the view temp are count and count.
Examples of the use of two other optional SQL statements which help to produce output in the correct form, GROUP BY and ORDER BY, are given in Box S4a and S4b. The GROUP BY statement in this example groups the occurrences of each unique attribute occurrence in the column county. ORDER BY count DESC ensures that the rows are given with the largest count first, i.e. in descending order. COUNT is useful either to count the total number of records of a particular type or, when used with DISTINCT, to find out the number of a repeating value, as in Box S4a. The DISTINCT statement removes duplicate rows in a column. The result of the query shown in Box S4b is given in Box O4b. Note that the view temp was issued in the same way as any other table.

The next query is slightly more complex as it relates information from five different tables in the database—hence two 'nested' select statements (see Box S9). The query is asking for the grid co-ordinates of a selection of sites to be output in a form that can be used to produce a distribution map—see Section 4.3 and Fig. 12. The sites selected are those from northern England which have samples from pit, well or ditch contexts, are dated to Roman or medieval periods and have a list of taxa which includes Agrostemma githago. Note that the statement which selects the plant name from the taxon_list does not find such things as Agrostemma cf. githago or cf. Agrostemma sp. This could be achieved by giving the statement: AND taxon_list CONTAINING 'Agrostemma' or AND taxon_list IN ('cf. Agrostemma githago', 'Agrostemma cf. githago', 'Agrostemma githago', ...)
Box 53: Select information about sites which have report numbers between 1000 and 1060 and are from grid squares SK and SP. Give the output in alphabetical order by site name.

```
SELECT sites.site, report, gridsq, county, period
FROM sites, reports
WHERE sites.site = reports.site
AND report >1000 AND <1060
(AND gridsq LIKE 'SK 43'
OR gridsq LIKE 'SP 42')
AND period NOT CONTAINING ?'
ORDER BY sites.site ASC;
```

Box 03: Alphabetical list of sites from selected grid squares

<table>
<thead>
<tr>
<th>SITE</th>
<th>REPORT</th>
<th>GRIDSQ</th>
<th>COUNTY</th>
<th>PERIOD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Austin Friars 73-8</td>
<td>1037</td>
<td>SK 43</td>
<td>LCS</td>
<td>pre-Boreal</td>
</tr>
<tr>
<td>Birmingham Moor 73-5</td>
<td>1042</td>
<td>SP 42</td>
<td>WND</td>
<td>early/post-med</td>
</tr>
<tr>
<td>Broom-on-the-hill 66</td>
<td>1020</td>
<td>SK 43</td>
<td>LCS</td>
<td>Iron Age</td>
</tr>
<tr>
<td>Fishbridge 74-5</td>
<td>1059</td>
<td>SK 43</td>
<td>WND</td>
<td>Iron Age/EB</td>
</tr>
<tr>
<td>Goodwin Cairns 53</td>
<td>1022</td>
<td>SK 43</td>
<td>DER</td>
<td>Beaker Culture</td>
</tr>
<tr>
<td>Harrold 76-8</td>
<td>1033</td>
<td>SP 42</td>
<td>WND</td>
<td>Bronze</td>
</tr>
<tr>
<td>Wellington 70-2</td>
<td>1025</td>
<td>SK 43</td>
<td>DER</td>
<td>Neolithic-Bronze</td>
</tr>
<tr>
<td>Wellington 70-2</td>
<td>1060</td>
<td>SK 43</td>
<td>DER</td>
<td>Neolithic-axial</td>
</tr>
</tbody>
</table>

Note that SQL processes the parts of the select statement in order, starting with those parts most deeply nested by brackets.

The string 'concatenation' operator (||), in Box 55, is used to link the east and north values with a comma: '11', '11'. It could also be used to insert blank spaces or a text string.

Finally in this brief description of SQL queries, SELECT statements can be used to check the integrity of the database. For example, in case a site has been recorded twice by mistake, the query in Box 56 finds site names which are duplicated in the sites table. In fact, as site is a primary key (see above Section 2) in the sites table there certainly should not be any duplicates, as the field should be given a PRIMARY KEY column constraint (see Section 24.8) when the table is created. This SELECT statement gives an example of the use of the HAVING predicate where SQL evaluates an example of the use of the HAVING predicate where SQL evaluates an intermediate result table which is then used in the next stage of the query.

The query in Box 57 finds if there are any site names which do not match between the sites table and the reports table. It does this by selecting first the ones that do match in the nested SELECT statement and then those not in the nested select.

4.3 Distribution maps

The grid references produced from a query such as the one in Box 05 can be fed into a distribution map plotting program or into a
Box 54a: Create a VIEW called temp with two columns county and count.

CREATE VIEW temp (county, count) AS
(SELECT DISTINCT county, COUNT(*)) FROM sites, reports WHERE sites.site = reports.site AND reports/report IN (SELECT report FROM reports WHERE type = 'C') GROUP BY county;

Box 54b: How many reports are there in each county (excluding reports which are only charcoal and wood identifications). List the twelve counties with the most reports.

SELECT counties.county, temp.count FROM temp, counties WHERE temp.county = counties.county ORDER BY count DESC LIMIT TO 12 ROWS;

classic geographic information system (GIS). A GIS can store complex geographic information in a database form and allows data to be analysed in relation to spatial variables stored as map overlays (see Waits 1988). Figures 11 and 12 were produced using ARC/INFO a GIS system on the mainframe computer at York University.

4.4 Summary

Once the arguably tedious process of developing a relational database, and collecting, loading and checking all the data has been completed, then the joys of using SQL can commence . . .

Acknowledgements

With many thanks to all my friends and colleagues who have waded critically through earlier versions of this text, adding constructive comments and helping to make this "ABC of SQL and the DBMS for the ARCD of GB and II"; all mistakes and opinions are the author's own, however. This work was funded by a Science and Engineering Research Council Instant Award which is gratefully acknowledged.

References


**Box S5:** Find occurrences of *Agrostemma githago* in selected contexts from Roman and medieval sites in the counties of northern England.

```sql
SELECT DISTINCT east || '°' || north
FROM sites, reports
WHERE sites.county IN ('NYR', 'LAN', 'DUR', 'NOT', 'LIN', 'CUM', 'CHE', 'WYR', 'HUM', 'DER', 'T&W', 'CLV', 'NHU')
AND sites.size = reports.size
AND report IN
(SELECT DISTINCT report
FROM reg_lists
WHERE list IN
(SELECT tList FROM contexts c, taxon_lists t
WHERE tList=c.list
AND c.period CONTAINING 'roman'
OR c.period CONTAINING 'medieval'
AND c.species CONTAINING 'pit'
OR c.species CONTAINING 'well'
OR c.species CONTAINING 'ditch'
AND c.taxon LIKE 'Agrostemma githago');
```

**Box S6:** Find all the duplicate site names in the sites table

```sql
SELECT *
FROM sites
GROUP BY site
HAVING COUNT(*) > 1;
```

**Box S7:** Find if any site names in the reports and sites tables do not match

```sql
SELECT site
FROM sites
WHERE site NOT IN
(SELECT reports.size
FROM reports
WHERE reports.size = sites.size);
```


**Appendix**

**Glossary of terms used in this paper**

**ASCII** a standard code used for representing characters

**atomic** values/attributes are atomic if they are nondecomposable (i.e., the smallest unit of data (a domain is a set of such values)

**attribute** a property of an entity. Entities have a defined set of attributes (cf. domain)

**attribute occurrence** atomic value occurring in each cell of a table row or tuple, the complete set of these forming a domain

**column/field** the vertical structure of a table

**constraints** help to maintain the integrity of the data by preventing the wrong type of data to be entered
Figure 11. Distribution map produced using ARQ-FNQ, showing all archaeological sites in the British Isles (in the ABCD) with 'in' type reports (see pp. 17-18 for explanation of report types).
Figure 12. Map showing sites where 'Agrostemma' has been recorded from pits, wells and ditches on Roman and medieval sites in N. England (cf. the SELECT statement in Box 55, but using the statement "IN.name CONTAINING 'Agrostemma'").

dependencies
domains
entity
identifier
index
null
primary key
relationship
schema
table
...
Short contributions

An eagle, *Haliaeetus albicilla* (L.), skull from Roman Leicester, England, with some speculations concerning the palaeoecology of the Soar valley

Ian L. Baxter

Leicestershire Archaeological Unit, Leicestershire Museums, Arts and Records Service, Museum House, 116 Hinckley Road, Leicester LE3 0RQ, U.K.

Summary

In 1991, during an evaluation excavation under the basements of shops fronting the High Street in Leicester, England, U.K., an eagle skull was found in a Roman soil. This has been identified as belonging to a white-tailed eagle, *Haliaeetus albicilla* (L.), together with bird remains from other sites in the city, there is a suggestion that rather different ecological conditions prevailed in that part of the Soar valley from prehistoric to medieval times.

Introduction

From November 1991 to January 1992, the Leicestershire Archaeological Unit carried out excavations in the basements of numbers 35 to 47 High Street, Leicester. The purpose of the excavation was to evaluate the surviving archaeological deposits in consideration of a planning application for the redevelopment of the site. This site was situated just within the east gate of the Roman town defences, 90 m north of the main east-west axis road, which was in fact the line of the Fosse Way as it passed through this major Roman town, *Rata Corieltaviis*.

Twelve trenches, with an average area of nine square metres, were opened up. Throughout the site, the construction of cellars in the 19th-20th centuries has caused destruction of the later archaeological levels and in this case approximately 2.5 m of these deposits had been lost. However, a considerable amount of archaeology has survived, mostly consisting of Roman and medieval cut features, predomi-

antly pits, dating from the 1st to the 13th centuries. In one of these, Feature F300 at 37 High Street, an eagle skull was found.

**Feature F300**

This consisted of a circular shaft, 1.8 m in diameter, with vertical sides (FIG. 13). It was excavated to a depth of 1.9 m, the maximum limit of the evaluation brief, but continued below this level with no indication of bottoming. In this basement, it was estimated that 1.3 m of Roman deposits above the excavation surface had been lost (J. Hagar, pers. comm.). Context 339 consisted of a clean brown friable silty sand and 333 may have contained evidence of a wattle lining. The other fills consisted predominantly of brown silts and in places there were traces of a 'cress'-like element. Some of these contexts produced finds of pottery, tile, animal bone, wall plaster, oyster shell and slag, but others were sterile. Feature F300 was probably a well and has been dated to the mid 3rd century (Lucas et al. forthcoming).

The animal bone in F300 was not typical of the site as a whole, containing a high proportion of wild species and immature individuals of domestic species. The following were notable: a foetal or neonatal piglet (Sus) in 314; fragments of a second piglet of similar age in 326; incomplete kitten (*Felis*?) skeletons from 316 (above 314) and 327; incomplete skeletons of young dogs (*Canis*, <18 months and 15 months) from 326 and 325 (below 316 and above 314); an incomplete crow (*Corvus corone* L.) skeleton from 337; and an incomplete jackdaw (*Corvus monedula* L.) skeleton from 315. Most of the domestic cattle (Bos) and sheep/goat (*Ovis/Capra*) fragments were from sub-adult animals (<18-30 months) and consisted most frequently of head and jaw elements.

The most unusual bone fragments, however, came from context 327 and 330, and consisted of a small goose *Anseranas* and an eagle skull, respectively. The goose has been tentatively identified as barnacle goose, *Branta leucopsis* (Bechstein), chiefly on the basis of size after comparison with reference material housed at Leicestershire Museums and the British Museum (Natural History) at Tring. This is a non-breeding visitor from September to May, mainly frequenting western Scotland and Ireland in recent times (Lucas et al., forthcoming). Barnacle goose has been found
in Roman contexts at Carlisle and York (Parker 1988, tables 1 and 2).

The eagle cranium

The eagle skull consists of a beakless adult cranium (Figs. 14-16) showing clear signs of having been decapitated in antiquity. Microscopic examination indicates that this was done with a very sharp blade, the blow coming from above and across the left occipital region (Fig. 16). There is also a chop mark on the left side above the processus frontalis which may represent an earlier failed blow. No trace of the beak was found during excavation or processing of the samples and there are no fresh breaks to suggest it was lost at that time; in all probability it had been removed prior to deposition. No other eagle bones were recovered from this or any other context within this feature or any other on the site and, indeed, it is the only eagle bone found so far in archaeological deposits in Leicestershire.

The skull has been identified as belonging to a white-tailed sea eagle, Haliaetus albicilla (L.), after comparison with reference material at the British Museum (Natural History). Tintin, on the following grounds:

(i) There is no fronto-parietal groove as in the golden eagle, Aquila chrysaetos (L.), but a shallow median groove on the parietal only, as in Haliaetus (Pycraft 1902, 280);
(ii) There is an occipito-parietal cranial nerve foramen, generally present in Haliaetus but not in Aquila;
(iii) The interorbital septum is not fenestrated. Out of eight white-tailed eagle skulls examined only one immature specimen (B.M. (N.H.) 5/1954.30.104) had a small fenestra in the interorbital, while only one (B.M. (N.H.) 1922.8.3.1) out of twelve golden eagle skulls was fenestrated. The latter is probably a very old specimen;
(iv) The breadth of the pars nasalis of the frontal (SB 110) is greater than Aquila specimens measured, but within the range of H. albicilla. The greater breadth of this region provides support for a more massive bill in white-tailed eagles (cf. bill measurements in Cramp 1980);
(v) Foramina in supraorbitals are more typical of white-tailed eagles than golden eagle;
(vi) The shape of the last temporal plate and paraphenoidal rostrum are similar to H. albicilla and differ from those found in Aquila (Fig. 15);
(vii) Contra Pycraft (1902, 280), the Eustachian channels in Aquila are not invariably open grooves (they seem to be partly ossified in the
Leicestershire Museums specimen B1952/146, for example, and those of Haliaeetus are not invariably ossified (they are open in B.M. [N.H.] specimens 1851.11.10.46 and 1862.3.30.2). They are open in the archaeological specimen from Leicester, which seems to be undamaged on the right side (Fig. 15).

While some of these morphological details may seem relatively trivial, others, such as (i), (iii), (iv) and (v) above, are more significant and combine to provide a reasonably secure identification. The interorbital septum and basi-temporal of the Leicester specimen are notably asymmetric (Figs. 14(B) and 15).

Archaeological eagles

The white-tailed eagle was present in Britain from towards the end of the penultimate (Wolstonian) glaciation until its extermination at the beginning of this century (Reid-Henry and Harrison 1988, 79). Their remains are more frequently found on archaeological sites than those of the golden eagle (Cohen and Serjeantson 1986, 2). The bones of around 14 white-tailed eagles were found in the chambered tomb of Ishister in Orkney (Bramwell 1983), and others have been found on prehistoric and historic sites in Wessex (Coy 1985). The remains of white-tailed eagle have been found at six other Roman sites in
**Table 1. Measurements of the eagle cranium from 37 High Street, Leicester** (based on von den Driesch 1976). LP—length from the protuberantia occipitalis externa to the most anterior points of the processus frontalis of the incisivum in the median plane; SBO—smallest breadth between the orbits on the dorsal side (the smallest breadth of the pars nasalis of the frontale); GB—greatest breadth along the processus postfrontalis. *Student's t* test: the difference is *not* significant for SBO or the two species greatly exceeds the tabulated value of *t* for 17 degrees of freedom and *p* = 0.001.

<table>
<thead>
<tr>
<th>Species</th>
<th>LP</th>
<th>SBO</th>
<th>GB</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Aquila chrysaetos</em></td>
<td>range</td>
<td>60.6-72.6</td>
<td>20.6-27.5</td>
</tr>
<tr>
<td></td>
<td>mean</td>
<td>69.24</td>
<td>24.76</td>
</tr>
<tr>
<td></td>
<td>S.D.</td>
<td>3.35</td>
<td>1.93</td>
</tr>
<tr>
<td><em>Haliaeetus albicilla</em></td>
<td>range</td>
<td>66.7-71.7</td>
<td>27.3-32.1</td>
</tr>
<tr>
<td></td>
<td>mean</td>
<td>69.27</td>
<td>29.67</td>
</tr>
<tr>
<td></td>
<td>S.D.</td>
<td>1.94</td>
<td>1.86</td>
</tr>
</tbody>
</table>

High Street, Leicester, 69.1

Britain: Droitwich, Uley, Duntisbourne, Sheepen (Cambridgeshire), and Southwark and Billinge- gate in London. Three of these are urban contexts and three rural, including a Romano- Celtic temple site (Parke 1988, 204, tables 1 and 2).

**Habits and habitat**

The white-tailed eagle is described by Reid-Henry and Harrison (1988, 78) as a 'lowland waterside eagle', while Brown (1978) considers them to have been coastal and estuarine in Britain. It is possible that in pre- and early historic times white-tailed eagles were more common inland than in the recent past. They do breed inland at the present day in Sweden, Germany and Finland (ibid., 86). Certainly, it would be hard to find a more inland locality in England than Leicester. Although young birds ringed in Norway have been found up to 60 miles from the place of ringing (ibid., 88), it would be exceptional to find an adult any great distance outside its home range. Generally a sedentary species with a home range of about 600-800 hectares, white-tailed eagles tend to be more sociable than other large birds of prey and may roost or feed in groups of five or six if food is abundant in a certain locality (ibid., 92). Frequently referred to as a companion of the wolf and raven at battlefields in Anglo-Saxon and Norse literature (Reid-Henry and Harrison 1988), even larger gatherings probably occurred after battles. Such gatherings happen opportunistically in present-day Alaska, where large numbers of a related species (bald eagle) flock during salmon spawning. Fish constitute about two-thirds of the normal prey of white-tailed eagles with diving birds, hares and carrion comprising most of the rest of the diet (Brown 1978, 87).

**Ecological implications**

The barnacle goose from 3rd century Leicester probably represents an individual blown off course, rather than being evidence of a regular visiting population. In recent years the Spitsbergen population winters on the Solway Firth, but the recovery in Norway of specimens ringed in the Netherlands suggests a possible origin of the rare but regular visitors to more southerly localities in England (Cramp 1977, 432).
Figure 15 (right). Ventral aspect of the eagle skull from 37 High Street, Leicester. p—parasphenoid, e—Eustachian channel, b—basitemporal.

Figure 16 (below). Occipital view of the eagle skull from 37 High Street, Leicester, showing area sliced through.
The generally coastal habitat of white-tailed eagles in north-western Europe (ibid., II, 49) suggests that the specimen under consideration may also have arrived in Leicester from the east coast. However, elsewhere white-tailed eagles frequent large rivers and inland waters. They may have bred at Windermere up to the 18th century (Brown 1978, 86). It is known that they show a preference for the proximity of colonies of wading and swimming species such as herons and cormorants. Both adults and eggs of these species are prey items, along with fish, diving birds and carrion (Cramp 1980, 49).

While there is a lack of data for the Roman period, it is known that the Soar (which runs through Leicester) is prehistoric (Shackley and Hunt 1985, 19), medieval (Mellor and Forrester 1981, 6) and in recent times was a sluggish river subject to extensive flooding, with the consequent creation of shallow ponds or lakes, marshland and reedbeds, not far distant from the Roman and medieval city. These circumstances persisted until the flood relief schemes of the late 19th century (Stevens 1972, 53).

Wader bones similar in size to snipe, Gallinago gallinago (L.), atern or redshank, Tringa totanus (L.), size and a spoon, Cygnus cygnus (L.), mandible have been found in Roman deposits at the nearby Shires site at Little Lane in Leicester, along with duck and goose bones that are probably domestic (Gidney 1991a). Further waterfowl remains were recovered from the medieval deposits at Little Lane and St Peter’s Lane (Gidney 1991b; 1991c). The wide range of wild bird species, including waders, in one post-medieval phase at St Peter’s Lane suggests wildfowling was practised at that time (Gidney 1991d).

Waterfowl account for nearly 50% of wild bird species recovered from the excavation of the medieval Augustinian friary, situated by the River Soar, about 500 m from the High Street site. These comprise mallard, Anas platyrhynchos (L.) (at least some of which may be domestic duck); cormorant, Phalacrocorax carbo (L.); grey heron, Ardea cinerea (L.); snipe, mégas abatis (L.) and crane, Grus grus (L.), (Thavley 1981, 173). All, with the probable exception of snipe, were native species until comparatively recent times and potential prey of the white-tailed eagle. Cormorant, heron, crane and snipe all nest in trees, as does the white-tailed eagle. The rare migrant snipe is said to like ‘drowned woodlands with many dead trees and overhows or other backwaters of large rivers’ (Cramp 1977, 669).

The Soar may be expected to have supported many fish. Analysis of the fish remains from the Shires site in Leicester, very near the High Street site, suggests that a greater proportion of freshwater fish was consumed by the Romans in that area at any time during the medieval era. Most of the freshwater fish were small, suggesting they were caught by net, but large examples of bleak, Leuciscus cephalus (L.), and tench, Tinca tinca (L.), were also represented (Nicholson, forthcoming). In 1307 the Dominicans were granted liberty to fish three day weekly in the river Soar with a net of convenient mesh so as not to destroy the young fish... and it has been suggested that the Augustinian friary in its early years may have obtained its fish exclusively from the adjacent river (Mellor and Forrester 1981, 15).

The presence of domestic livestock, particularly sheep and goats, in the farmyards and villas outside the Roman city would have provided a regular source of carcase, especially during lambing and in winter. Urban butchery sites such as slaughteryards would have provided similar opportunities, which ravens, kites and buzzards are known to have exploited in Leicester and elsewhere.

The white-tailed eagle may thus have been resident in the Soar valley or a visitor from the coast. Given the probable environment at the time and the sedentary habits of the species, an adult bird such as this was probably a resident. It may have been killed as vermin (white-tailed eagles were exterminated in 19th and early 20th century Britain as sheep stealers) or to provide feathers for fetiching. The absence of any post-crannial remains and of the bill leaves more room for less precise speculation, however. The beaks and claws of eagles and other raptors seem to have been used as talismans in prehistoric Europe (Jack 1948, 129-30), and it is well known that the Romans associated the eagle and the thunderbolt with Jupiter. A neolithic polished stone axe-head (a thunderbolt?) was found in an earlier Roman pit on the same site only 15 m away from that containing the eagle skull (Lucas et al., forthcoming). The present evidence is too slight, however, to provide an adequate basis for cultural speculation and the presence of the eagle skull in the backfill of a Roman well must remain problematic.

36
Acknowledgements

I would like to thank Jennie Coy, Jan Dawson of Leicestershire Museums and Jo Bailey of the British Museum (Natural History) at Tring, for arranging access to reference material in their charge; David Hopkins for his drawing of the eagle skull; Julian Hager, Patrick Clay, Angela Monckton and John Lucas at the Leicestershire Archaeological Unit for their contextual information and support in the writing of this paper; and Terry O'Connor of Bradford University for checking my identification of the goose bone. Tony Coulthould or Leicester University for help with references, and both for their encouragement.

References


Nicholson, R. A. (forthcoming). Fish remains from excavations at The Shires: Little Lane (A39) and St Peter's Lane (A40), Leicester. Ancient Monuments Laboratory Report.


Disk copy received August 1992, accepted October 1992

Papers from the bone taphonomy workshop at York, September 1991

An investigation into the effects on fish bone of passage through the human gut: some experiments and comparisons with archaeological material

Rebecca A. Nicholson

Environmental Archaeology Unit, Department of Biology, University of York, York YO1 5DD, U.K. (address for correspondence: 'Stans Hall', 34 Moorside Road, Eccleshill, Bradford BD2 5EL, U.K.)

Summary

The survival rate and appearance of chalcedony fish bones after consumption by a human is examined, and comparisons made with assemblages of chalcedony bones recovered archaeologically. Several sizes of fish were ingested, on five separate occasions. While the proportion of bones surviving in a recognizable form varied, the types of skeletal elements surviving remained fairly constant. In all cases the proportions of bones surviving was very low, but the relationship between fish bone survival and fish size was not straightforward. While all bones from very small fish were lost, those from the largest fish survived well. Only bones from smaller individuals. Characteristic forms of crushing and etching were observed, and these forms of damage, as well as the skeletal element distribution, are used to try to distinguish excavated assemblages of bones originating in human faeces from fish bone originating from other sources.

Introduction

Background

In recent years there have been a number of investigations into the means by which different predators may be recognised from the types of damage they inflict on bones. These studies have principally concentrated on large carnivores (Brain 1981; Bischof 1981; Stallibrass 1986; 1990) and avian raptores (Mayhew 1977; Dodson and Wexlar 1979; Korth 1979; Denys 1985; Andrews 1990) as predators on large and small mammals. However, for most archaeological sites, the principal accumulator of bones was probably man. At least a proportion of the smaller bones are likely to have been ingested rather than discarded before eating, and each bone would be predicted to be recovered from archaeological sites in considerable numbers.

Occasionally coprolites and pellets are recovered intact from archaeological excavations, usually from dry sites but sometimes in desiccated or mineralised form (e.g. the human coprolite from the site of 6-8 Pavement, York (Jones 1983)). Many of these appear to have been dog coprolites (e.g. Dinahsey 1966; Paap 1976; Jones 1990). Other archaeological studies of intact ancient human coprolites containing bones include reports of some containing numerous fish remains, for example those from Lovelock Cave, Churchill County, Nevada (Bollett 1967; 1970). Intact human as well as dog coprolites have also been recovered from Lake Calakulla, Coacchella Valley, California (Wilke 1978). However, it is fairly unusual for whole coprolites to be identifiable in archaeology. More often the matrix will disintegrate, making their contents less easy to recognise. It is for this reason that investigations into criteria by which digested remains may be identified are required.

Despite the prevalence of contexts interpreted as possible cess pits during excavation, and the likelihood of a human faecal component in many, particularly urban, archaeological deposits, there have been relatively few studies of the components of modern human faeces. Considering the unpleasantness of the task, this is not altogether surprising. Examining coprolites from antiquity is much less romantic, and more socially acceptable, than studying contemporary faeces.

There is a very small literature detailing experimental work into the effects on organic materials of passage through the human gut. In 1977 Calder published one such account. Among the organic materials tested were scales of aounder (Rhomboscalea sp.) and a seel (Peltotornipes novae-zelandiae). The shark dermal denticles (species unspecified), limpets (Calyptraeidae) radulae and periwinkle (Littorinidae) opercula. Of these, the scales were completely digested, while almost all of
<table>
<thead>
<tr>
<th>Experiment No:</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Species ingested:</td>
<td>kipper</td>
<td>sardines</td>
<td>sardines</td>
<td>sardines</td>
<td>sardines</td>
</tr>
<tr>
<td>No.</td>
<td>1</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>*Approx. no. identit</td>
<td>80</td>
<td>80</td>
<td>80</td>
<td>80</td>
<td>80</td>
</tr>
<tr>
<td>bones in 1 fish</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. whole vertebral con</td>
<td>1</td>
<td>18</td>
<td>14</td>
<td>4</td>
<td>8</td>
</tr>
<tr>
<td>No. other bones</td>
<td>1</td>
<td>6</td>
<td>4</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>recovered</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. eye lenses</td>
<td>0</td>
<td>1</td>
<td>7</td>
<td>8</td>
<td>6</td>
</tr>
<tr>
<td>recovered</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. otoliths</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>recovered</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>% survival</td>
<td>2.5</td>
<td>6.0</td>
<td>4.5</td>
<td>1.3</td>
<td>2.3</td>
</tr>
<tr>
<td>whole fish</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>% survival of</td>
<td>1.8</td>
<td>6.7</td>
<td>5.2</td>
<td>1.5</td>
<td>3.0</td>
</tr>
<tr>
<td>all bones</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>whole vertebral</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>con</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2. Survival of fish bones through the human digestive system (excluding the 25 small herring/sprats eaten with the kipper in Experiment 1, from which no identifiable bones survived). * The expected numbers of bones in one fish (80) is based on the figures given by Jones (Wheeler and Jones 1989) to facilitate comparison with Jones' experiments. However, Jones does not detail the elements included in this figure for herring, so that the exact skeletal elements considered by him are not known. The present author considers a much greater number of head and pelvic elements (bones of the cranium, skull and pelvic region excluding spines, ribs, rays and branchial bones, etc.) to be identifiable. Otoliths and lenses are not included in the 'bone' category.

the mollusc parts and shark denticles were recovered intact.

Jones (1984; 1986) examined the loss of bone after passage through the digestive tracts of a dog, a pig, a rat and a man, and documented the types of damage seen on the very few fragments which survived the process. Payne and Manson extracted bones after fish were fed to a dog, and these were examined by Jones (details in Jones 1984).

Objectives and approach

Undeterred by the social consequences, this author set out to augment Andrew Jones' courageous work and further investigate the effects on fish bone of human digestion. Jones himself ate only one kipper, so there was clearly a need for replication to examine the extent of variation in patterns of bone survival. The principal objective of the experiments described below was therefore to investigate whether the sorts of bones which
survive the human digestive process, and the types of damage inflicted on them, is predictable, or whether variability in digestive efficiency leads to great variation in the resulting bone assemblage. From an archaeological viewpoint, the experiment was therefore designed to assess the extent to which humans modify bone assemblages during ingestion and digestion and, secondly, to assess the potential for recognising in archaeozoological assemblages bones which have passed through the human digestive tract.

To look at the effects of human digestion on ingested fish bones a number of complete fish were eaten. Several sizes of fish were used, the limits to size and species being set by the feasibility of swallowing the bones. The digested bone assemblages were then compared with assemblages of small fish remains recovered from archaeological deposits, in an attempt to determine whether the archaeological bones had, in fact, been deposited in human faeces. Human-digested bone assemblages are elsewhere compared with assemblages of small fish bones recovered from otter spraints, seal scats, gull pellets and water-abraded skeletons (Nicholson 1991a).

Materials and methods

Whole fish, lightly cooked, were eaten by the author on five separate occasions. On the first occasion one kipper (Chlops bairdii L. total length 300 mm) and 25 whitebait (young herrings, also C. bairdii L. and sprats Sprattus sprattus (L.), of lengths from 60 mm to 80 mm) were eaten. On the subsequent occasions five sardines (Sardina pilchardus (Walbaum) of total lengths from 160 mm to 190 mm) were consumed. Each fish was eaten in its entirety, after frying or grilling for up to five minutes; this caused charring to the fins,
but otherwise the bones appeared undamaged. With each fish meal, approximately 200 g of tinned sweetcorn (Zea mays L.) was eaten, which acted as a marker to indicate when the entire meal had passed through the gut, as recommended by Calder (1977). Bread was also consumed, as the bones (especially the kipper’s head bones) were sometimes difficult to swallow. It was found that the fish heads required more mastication than the vertebrae. Many vertebrae were probably swallowed unchewed. Faeces were collected for five or six days after each sample of fish had been eaten. These faeces were soaked in warm water for up to 24 hours. Disaggregation proved to be possible without recourse to the chemicals described by Calder (ibid.), by passing a stream of hot water over the faeces, held in a 500 micron mesh sieve. The residue was further cleaned by moving the base of the sieve up and down in a shallow bowl of warm water. The residues were sorted either wet or after drying at 40°C. Bones were picked out using a low-powered dissecting microscope.

Results and discussion

Of all the complete fish eaten, very few bones survived the digestive process. Details of the recovered fragments are given in the appendix and summarised in Table 2. Figure 17 illustrates graphically the proportional representation of skeletal elements recovered, based on the pooled results from all the fish ingested.

The proportional representation of skeletal elements (PR) has been calculated using the method given by Dodson and Wexlar (1978, table 1) and is based on the numbers of bones surviving compared with the expected number:

$$\text{PR} = \frac{F_o}{F_e \times \text{MIN}} \times 100$$

where $F_o$ = the number of recovered bones, $F_e$ = the expected number of the element in one individual.

The minimum number of individuals (MIN) is calculated in the conventional way, from the most commonly represented bone in the assemblage. It should be noted that this figure is a minimum; the most commonly represented bone will probably have suffered some loss too.

The extremely low numbers of bones which survived digestion are of similar proportions to those reported by Jones (1986) and Wheeler and Jones (1988, 73-4). Most bones were damaged and, of those identified, many would not have been identified to species had I not known what was swallowed. None of the whitebait bones survived in any form. Figure 17 illustrates the extent of bone loss, expressed as the proportional representation...
Figure 20. Proportional representation of clupeid skeletal elements from Viborg Sønderse, Denmark.

Figure 21. Proportional representation of clupeid skeletal elements from Redcastle Farze, Thetford, England.
Figure 22. Proportional representation of clapoid skeletal elements from Queen Street, Newcastle-upon-Tyne, England.

Figure 23. Proportional representation of clapoid skeletal elements from Crown Court, Newcastle-upon-Tyne, England.
of skeletal elements; thus the recovered assemblage is treated as if it were from an archaeological sample in which the number of individuals would be calculated as a minimum (MNI) rather than known in advance. The range of potentially identifiable skeletal elements listed in Figure 17 is based upon those bones which this author considers identifiable among the Clupeidae, and includes a greater number of skeletal elements than considered by Jones (see Table 2). Brevity and the small number of bones which survived, and the relatively low numbers of fish ingested, only very crude conclusions may be drawn about the relative survival rates of different parts of the skeleton, by calculating the percentage rates for vertebrae and for the entire skeleton as a unit (Table 2).

Of all the parts of the skeleton, for the species used in these experiments, the eye lens was by far the most resilient element. Unfortunately, as eye lenses are not calcified their survival in all but exceptional sediments is unlikely. Lenses might also be misidentified as seeds. Thus concentrations of eye lenses cannot be used as a diagnostic feature of deposits containing human faecal material.

The destruction of almost all head bones indicates that it may not be possible to say whether complete fish were consumed from assemblages of bones contained in faecal material. The lower survival rate for the larger kipper (herring) when compared with the sardines may result from the greater amount of mastication necessary to enable swallowing of the kipper bones. As all whitebait bones were lost, chewing clearly does not provide the whole explanation for differences in survival, however. Bone size must also play a part, as may the extent of calcification, which will be related both to fish species and age. The survival of scale fragments is contrary to the results given by Calder (1977) who found that no fish scales survived digestion.

Most of the vertebrae recovered were crushed; the most common form of damage was compression and breaking of the spinous processes of the two articulating facets (Fig. 18). This commonly resulted in the two halves being separated. Many fragments of the centrum rim were also recovered. Several vertebrae were stained black or dark brown. These specimens commonly showed extreme acid dissolution, causing crenellation and rounding to the edges of the articulating facets (Fig. 19). Erosion causing smooth crenellated edges was also observed on other fragments, not all of them identifiable. There was no obvious correlation between extent of bone erosion and the length of time before the bone was passed in the faeces. Complete, unstrained and unbroken vertebrae were recovered with stained and eroded specimens.

Evidently the efficiency of the digestive system varies, even in one individual. As only fish of the Clupeidae were consumed in these experiments it would be useful to use fish of other families and to compare the results. Eel, Anguilla anguilla (L.), and stickleback, Gasterosteus aculeatus (L.), were commonly found in deposits where human faeces are suspected, and other small fish could also potentially be consumed whole. It is to be expected that the efficiency of the digestive system will vary between individuals and depending on the health of the individual. Other variable factors must also include the extent of chewing, itself influenced by the number and condition of teeth, and the amount of "paddling" surrounding the bone when consumed. Preparation techniques, such as boiling, roasting and drying may also affect the extent to which bones will be digested. Archaeological deposits containing faeces are frequently identified by the presence of abundant intestinal parasite eggs (see below), and heavily parasitised individuals may digest food in a different way from healthy individuals. Stomach upsets would also be likely to reduce the efficiency of digestion. The experiment described here only provides a starting point; ideally many more experiments should be undertaken, using a number of different individuals, to establish the variation in digestive function.

Do different mammalian digestive systems affect bone differently?

Inevitably, in the absence of extensive experiment replication a cautious approach must be taken when attempting to identify the consumer from an assemblage of apparently chewed and digested bone. Considering the experiments which have been performed by this author (Nicholson 1991a) and by Jones (1984; 1986; and Wheeler and Jones 1999, 70-2), however, certain traits may be suggested.

When small (120 mm) complete herring and small complete place (Pleuronectes platessa L., 130 mm) were offered to rats Jones found that
all bones were completely ingested; none were present in the faeces. A larger herring (270 mm) was also offered to the rats, and some remains were not removed from the cage to prevent excessive decomposition and smell. Jones fed all these bones would have been ingested, too. A pig eagerly ate a large fresh herring (325 mm) and mackerel (Scomber scombrus L., 355 mm), producing only seven identifiable fragments per fish in its faeces. No descriptions of damage to the bones were given. A dog consumed a large herring (273 mm) and haddock (Gadus morhua L., 325 mm), lightly fried. Nine identifiable herring bones and ten haddock bones were passed in the faeces, most exhibiting damage from chewing.

Etching of the centroid facet of one herring vertebra is illustrated in Wheeler and Jones (1989, 73) and is similar to that observed by this author on some human-digested vertebrae. Two red snappers (Lutjanus campechanus) of about 450 g were fed to a dog by Payne and Munson (Jones 1984). Even for this relatively robustly-boned fish over 80% of the ingested bones were lost and, of those surviving, most were chewed and broken, and some etched and corroded. Of all the elements, the eye lens survived best, as also found by this author after human digestion. To summarise this evidence, given the small numbers of bones surviving digestion in pig, dog and swan, and the extensive damage inflicted to most bones on ingestion by all these animals, it seems unlikely that in the absence of other evidence (such as associated parasite eggs) the mammal responsible for small numbers of chewed and etched fish bones can be identified with any degree of certainty.

By contrast, this author has also examined fish bones from contemporary otter (Lutra lutra (L.)) rip-ups from Scotland and Shetland (details in Nicholson 1991a), and found remarkably little damage to the fish skeleton. Although the number and species of fish consumed was not controlled, it was clear that as far as small fish (<under 150 mm) are concerned almost no discernible loss of bone occurs with digestion. All bones of the skeleton were present, including otoliths, and signs of chewing and etching were restricted to a relatively small proportion of bones (often less than 30% of vertebrae appeared to have been chewed). Bones from larger fish were generally much more heavily chewed, and usually at most only one or two of these larger bones were present in a single spraint.

Presumably the difference relates to the otter’s ability to swallow small fish in their entirety, with minimal mastication, while large fish are treated with more care. Bones from grey seal (Halichoerus Grypus Fabricius) droppings, also examined by the author (Nicholson 1991a), frequently displayed signs of corrosion, but again otoliths seem to survive particularly well, even when from small fish such as sand eels (Ammodytidae).

Archaeological material

Having established that fish bones sustain considerable loss and damage during chewing and passage through the human gut, archaeo-

logical material from supposed cess pits from two sites—Viborg Sandervor and Thetford Redcastle Furze—was examined to see whether the fish bone exhibited similar features.

(a) Viborg Sandervor

Fish bones were recovered from waterlogged richly organic layers with a faecal component, excavated from beneath structures of Viking Age date at Viborg Sandervor, Denmark (D. E. Robinson, pers. comm.). The bones were recovered from two soil samples of wet weights 438 g and 375 g. These samples were sieved to 500 microns and the residues sorted using a low-powered binocular microscope. Further details are given in Nicholson 1991b.

In total, 201 bones were examined from context 1677, and 143 bones from context 1662. The faecal component in these deposits was inferred from the botanical remains and from the presence of eggs from the whipworm Trichuris trichiura (L.), a human gut parasite (Robinson and Boldsen 1989).

The fish species represented included eel Anguilla anguilla, herring Clupea harengus and possibly sprat Sprattus sprattus, flounder Platichthys flesus (L.), perch Perca fluviatilis L., bloat Alburnus alburnus (L.), possibly dace Leuciscus leuciscus (L.) and other cyprinids not further identified. With the exception of some of the herring bones (from individuals between 200–350 mm length) and the flounder dermal denticles (representing fish of at least 400 mm long), all bones were from fish of less than 150 mm total length. (Total length was estimated by comparison with bones from modern fish of known length.)
A small number of bones appeared charred, as indicated by a black peeling layer seen under the light microscope and confirmed under the scanning electron microscope. Only slight charring was observed, consistent with burning during cooking rather than resulting from rubbish disposal. Both assemblages also included a proportion of vertebral centra distorted in a manner consistent with chewing prior to swallowing, but most bones were complete (Table 3). The number of bones was too small to draw firm conclusions about the representation of skeletal elements; however, it is worth noting that a very limited range of bones was present. No bones appeared to have been partly dissolved or etched. Complete centra included a number from very small (under 100 mm total length) herring or sprat. The survival of these bones through the gut seems unlikely given the results of the experiments detailed above. No bones from the small clupeids (whitebait) survived digestion in the experiment although, as discussed above, it might be postulated that digestive disorders could affect the extent of bone destruction.

Without many more experiments it is not possible to state with certainty that undistorted, complete tiny clupeid vertebrae could not withstand digestion, but the results of these experiments indicate that it is unlikely, it appears, therefore, that while a faecal component was present, many of the fish remains from Viborg Sønderse probabilis probably represent table waste, and/or fish discarded at the food preparation stage—possibly from the guts of larger fish.

(b) Thetford Redcastle Fuze

Fish bones were also recovered from excavations at Thetford Redcastle Fuze, Norfolk, England. The bones had been recovered from samples of sediments which had been wet-sieved through 1 mm mesh (further details in Nicholson, forthcoming).

The assemblages considered here were recovered from five contexts: 795 (dated to the early Saxo period), 588 and 1677 (late Saxo), 1719 and 1720 (medieval, probably fourteenth century). All but the first were considered, possibly to have been censuses by the excavator (P. Murphy pers. comm.), and bones from context 965 were covered in a calcareous concretion presumed to be mineralised faeces.

Table 3. Fish remains from Viborg Sønderse, contexts 1677 and 1682.
Table 4. The fish remains from Redcastle Farze, contexts 795, 565, 1677, 1719, and 1720.

Species included eel Anguilla anguilla, herring Clupea harengus, Cypseloidae (probably small herring or sprat), stickleback Gasterosteus aculeatus, Cyprinidae (not further identified), and possibly ruffe Gymnocephalus cernuus L. and perch Perca fluviatilis. Apart from the herring and eel bones, the bones were from fish of under 150 mm total length. Table 4 details the bones recovered, and the numbers of charred and crushed or chewed bones. Each of these contexts contained a proportion of bones crushed in a manner consistent with the damage observed after passage through the human gut. No bones appeared to have been etched in the way seen in some of the bones from the digestion experiments detailed above.

However, a very limited range of samples from contexts 1719 and 1720 with herring bones were also examined for parasite eggs. Both deposits were found to contain eggs of the human whipworm Trichuris trichiura (Nicholson, forthcoming), supporting the conclusion that these contexts were cesspit fills, and that therefore at least a proportion of the fish bones had passed through the human gut. The similarity in the condition of the fish bones from the other contexts listed above indicates that these too had originated in cesspits. Bones from other contexts at Redcastle Farze did not appear crushed, although similar species and sizes of fish were represented.

47
Discussion

As the experiments into human digestion have only concerned fish of the Clupeidae, comparisons of body part abundance can only properly be made with similar species recovered archaeologically.

In both the Viborg and Thetford assemblages, fish remains from contexts not interpreted as containing faeces were too, so no comparisons however, between the two components of probable cespit and non-cespit origin could be made. The relatively low numbers of clupeid remains recovered also limits the possibilities for comparison between the archaeological and experimental material.

Herring remains are, however, frequently recovered from archaeological deposits where they do not appear to have arrived in faeces. Two sites which produced herring remains for which an origin in "ces" was not suspected are the sites of Queen's Street and Crown Court, Newcastle-upon-Tyne, England. Clupeid bones were recovered by wet sieving large (usually 60 l) samples through a 1 mm or a 0.5 mm mesh. The deposits were of mixed origin, but most of the herring remains from Queen Street were recovered from waterlogged deposits interpreted as medieval urban refuse tips. Further details of the deposits and fish bone from them are given by Nicholson (1988; 1989).

Statistical comparisons were not feasible owing to the low numbers of clupeid bones recovered during the experiments and from Viborg and Thetford. The proportional re-

representation of skeletal elements is illustrated by Figs. 20-3 for all the archaeological assemblages, for which the results from all contexts have been pooled.

Perhaps the clearest difference between the Newcastle assemblages when compared with the other bone groups is in the lack of chewed bones from the former. This in itself suggests that the Newcastle bones were not deposited in faeces. Additionally, the recovery of a very restricted range of bones from the material from Viborg and Thetford was reminiscent of the results from the experimental series, and was not in keeping with the trend observed for the Newcastle assemblages. The Viborg and Thetford material included a much greater proportion of vertebrae than in the experimental assemblage, however. This trend is apparent even when the small clupeid vertebrae from Viborg were excluded (as they may not have been consumed, see above). Several explanations are possible. The heads may have been discarded elsewhere, if fish were beheaded for the table, for example. Alternatively, crushed head bones may have been recognised from the experimental material (as all bone remains must have come from the ingested fish), but not from the archaeological material (where small, crushed bones may have been overlooked or classed as undigestible). It is also possible that surviving head bones were preferentially destroyed after burial. As head bones from herring are thin and fragile, their destruction in preference to vertebrae is likely (as shown in further experiments detailed in Nicholson 1991a), and chewed bones would probably be more at risk than complete bones. Given the small sample sizes further interpretation is unjustifiable; it is the gross differences between the Newcastle assemblages when compared with the other groups (from Viborg, Thetford and the experimental series) which provide the clearest evidence for different depositional origins. The presence of parasite eggs from the human whipworm indicates that in the case of the bones from Thetford and Viborg the consumer was probably Homo sapiens.

Conclusion

This study has confirmed the very extensive loss of bone which occurs due to passage through the human digestive system. From a whole ingested fish of size around 400 mm, only around 5% of bones survived. For very small fish these experiments have demonstrated that it is likely that no bones will survive the digestive process. After digestion a proportion of the bones show characteristic etching, crushing, and staining which may be recognised on archaeologically recovered specimens. This, as well as the very limited range of skeletal elements which appear to withstand passage through the gut of most mammals, provides good evidence for a faecal origin for bone assemblages. Determining the consumer is a more difficult task, however, as other mammals also damage bones to a similar degree. As far as possible caution must be used, and other evidence, such as the presence of gut parasite eggs, sought. In certain cases the size of the ingested bones may provide a clue; common sense would suggest that humans are unlikely to have swallowed very large bones. Considering the
number of small fish bones likely to have been lost as a result of human digestion, interpretations involving quantification or estimation of meat weights which are drawn from assemblages of bones from contexts where a 'cess' component is suspected should not be attempted.

Acknowledgements

Many thanks to Allen Hall and Harry Kenward and to an anonymous referee for their constructive comments on this paper, and to Borne Jones for inspiration!

References


Appendix

Fish remains recovered after ingestion by a human

Experiment 2

Fish ingested: 5 sardines, total lengths 170-185 mm.

Day 1 1 crushed vertebral centrum, 2 scale fragments

Day 2 1 eye lens, 2 chitinous fragments, 1 prey-opercular fragment, 1 acroporal, 1 stained hyomandibular fragment, 2 palatines (1 acid etched), 1 complete vertebral centrum, 8 laterally crushed vertebral centra, 1 acid etched and stained vertebrae, 11 vertebral fragments, 11 scale fragments, 2 branchiostegal rays, about 30 unidentified fragments

Day 3 4 crushed, stained and etched vertebral centra, 8 medially-laterally crushed, etched and stained vertebrae, 12 etched and stained vertebral fragments, 4 scale fragments, 20 unidentified fragments

Day 4 20 unidentified fragments

Day 5 No remains.

Experiment 3

Fish ingested: 5 sardines, total lengths 170-190 mm.

Day 1 1 eye lens, 1 complete vertebral centrum, 2 laterally crushed vertebrae, 1 vertebral fragments, 2 unidentified fragments

Day 2 5 eye lenses, 2 crushed epiphysyal, 1 crushed otic bulla, 1 stained articulae, 1 chitinous fragment, 2 complete vertebral centra, 1 etched and stained vertebrae, 16 vertebral fragments, 6 scale fragments, about 50 unidentified fragments

Day 3 1 eye lens, 1 complete, crushed vertebral centrum, 2 acid eroded and stained vertebral centra, 3 vertebral fragments, 3 scale fragments, 17 unidentified fragments

Day 4 1 etched vertebral centra fragment

Day 5 No remains.

Experiment 4

Fish ingested: 5 sardines, total lengths 180-190 mm.

Day 1 2 crushed vertebral centra, 2 vertebral fragments, 1 scale, 3 unidentified fragments

Day 2 2 eye lenses, 1 crushed vertebral centrum, 1 etched and stained vertebral centrum, 2 vertebral fragments, 1 etched epiphysyal, 1 otic bulla fragment

Day 3 4 eye lenses, 1 vertebral fragment, 3 unidentified fragments

Day 4 2 eye lenses, 1 unidentified fragment

Day 5 No remains.
Experiment 5
Fish ingested 5 sardines, total lengths 170-180 mm.

Day 1
2 crushed vertebral centra, 1 complete vertebral centrum, 3 vertebral fragments, 3 scale fragments, 2 unidentified fragments

Day 2
4 crushed vertebral centra, 2 complete vertebral centra, 4 eye lenses, 16 vertebral fragments, 1 horn hyaline, 10 scale fragments, about 50 unidentified fragments

Day 3
1 crushed, etched and stained vertebra, 5 vertebral frags, 2 eye lenses, 7 unidentified fragments

Day 4
No remains.

Day 5
No remains.

Final disk version received September 1992

[Editors' note: we apologise to the author for the delay in publication of this contribution, which was only in part beyond their control]

Notes, enquiries and correspondence

To the editors of Circum:

Congratulations to James Greig on his well-researched etymological discussion on what to call organic pit-fills (Circum 8, 70-3). Unfortunately we are often stuck with terms that have no basis and I would plump for *cappit*. I still cannot bring myself to speak or write his other term *s-at*, which is not surprising when one considers that my parents were brought up in Edwardian times and my grandparents in the Victorian period.

But I have found that etymological aspects can throw important light on archaeological and historical studies, and Allan Hauf's aside (ibid., 73) is an example. I looked up *shire tretise* from hemp or flax in my Dictionary of Textile Terms and found it listed as *spitens* "woollen trade term for all vegetable matter found entangled in wool"—an interesting example of the way in which a meaning can change with the context. This 524-page dictionary published by the now defunct Textile Mercury has no date and no pretension to be other than a glossary, but it cannot have been published later than about 1950 and is fast becoming for me an encyclopaedia of textile history.

Michael Ryder, 4 Osprey Close, Landissed, Southampon, SO1 3EX, U.K.

Book notices


This book is a remarkable monograph dealing with all aspects of *latis tinctoria*—including the botany, cultivation, and multifarious uses of the plant, the extraction of the famous blue dye, and the economic importance of woad in Europe in the Middle Ages. The first edition (of 1930) has now been faithfully reprinted in the U.S., giving it a new lease of life, the only (small) failure in reproduction being to lose the colour originally used in Plates I and XVI.

The original title page bore the attribution 'by the Late Jameson B. Hurry'; the author, sadly, died before seeing the completion of his work. He was a evidently a remarkable man for, as the memoir by a friend, Warrer R. Dawson, records, Hurry was by profession a general practitioner, having trained as an obstetric physician and then as a ship's surgeon. His interests were wide—something almost de rigueur in a professional man born in the middle of the last century—and one aspect of this can be illustrated by reference to a series of monographs he wrote on *Vicious Circles*—*Vicious Circles in Disease,...of Neurotheria,... and... in Sociology, then Poverty and its Vicious Circles*, not to mention his *Nursing on a Provident Basis*. Another series of publications relates to Reading Abbey and show his interest in history, and he combined history and medicine in works such as *Imbriob, the Visor and Physician of King Zoser and afterwards the Egyptian God of Medicine*. At his home in Reading he established an 'educational garden', bringing together economic plants from all over the world, making this and a museum of plant products freely available to visitors. And it was his interest in economic botany which 'found its last expression in the admirable monograph of the history and use of Woad'.

51
Perhaps the greatest value of this book—not surprising, given the eclectic interest of its author—is the breadth of its scope. On pages 252-5, for example, we find that wood was used for the treatment of a great range of diseases and conditions (according to authorities from Pliny and Galen to L'Obel and Parkinson), including erysipelas, angina, cancer and homoeopathic whilst on page 268 measures taken in Germany in the 17th century to prevent the use of imported indigo from Asia in preference to home-grown wood are detailed, all this with footnote references and additional notes where appropriate.

For British readers, however, perhaps the most intriguing part is the discussion of the growing of wood, mainly in the Fens of S. Lincolnshire, it is the latter part of the last century and the first two decades of the present one. Particularly fascinating are the monochrome photographs of structures and activities associated with wood-growing and manufacture. There is one, for example, of a wood mill at Parson Drove near Wisbech (pl. iv), with sketches of the interior (with wooden rollers for crushing the leaves) and walls (made of turf blocks—these were often short-lived structures made by the woodmen or waddies who moved from area to area as the wood crop exhausted the soil in one place).

Plate XIV was especially poignant for me, as it shows a German wood mill at the village of Hedingsieben, near Gotha, and a wood hall of the 16th century in Gotha itself. Both of these structures have survived—I was fortunate to be able to visit them both during the International Wood Conference held in Erfurt in June 1992. The caption to the photograph of the mill reminds that it ‘was pulled down a few years ago’, but it has been reconstructed and consists of the upright and horizontal stones and new wooden posts linking them.

Of interest to those of us engaged in the debate as to whether it was wood with which the Ancient Britons painted themselves blue, is Hurry’s description of ‘primitive methods of extracting indicars’ [the blue pigment]. He cites Plowright, writing at the beginning of the 20th century, who offers several suggestions as to how such skin colouring might have proceeded. It must be said, though, that wood varies enormously and rather unpredictably in its content of indicar (or indigo as it is more commonly called today); even with the suite of chemicals available to us, with which to effect the necessary reduction of indigo to a soluble form (which can then be oxidized into the familiar blue pigment). Dyeing with wood seems to be very much an art rather than a science!

Allan Hall, Environmental Archaeology Unit, University of York, Heslington, York YO1 5DD, U.K.

Editorial

Despite our best intentions, we have not managed to make up the whole of the backlog of issues of Circaus. However, we have made some progress, and we have enough material for the second part for 1992 and a good proportion of what we need for the first part of volume 11 to hand as we publish this issue.

We still need material, though! We remind potential contributors of the existence of the Notes, enquiries and correspondence section for very short pieces of copy.

Circaus is more and more taking on the character of a ‘proper’ journal, so that the somewhat informal Editorials we have indulged in until now are perhaps no longer appropriate. This will probably be the last.

In keeping with the more formal nature of Circaus, we intend to produce title pages and lists of contents for binding into volumes in future. If any readers have views on the format of this—e.g. do we include contents arranged by author and subject?—they should let us know. It would be impracticable for us to produce a full index with current resources.
Notes for Contributors

Articles for Circsea should be typed double-spaced on A4 paper with generous margins. Line drawings should be in black ink on white paper or drawing film, to fit within a frame 153 x 250 mm maximum. Captions should be supplied on a separate sheet of paper, and labelling on figures should either be in 'Learasit' (or an equivalent) or should be in soft pencil. Halftone photographs can be accommodated, but authors wishing to make extensive use of photographs, or colour, should note that they may be asked to contribute towards the high cost of production. The editors will modify short contributions to fit the layout and convention of Circsea. The same principle will be applied to idiosyncrasies of spelling and punctuation. Scientific articles will be submitted to referees; authors may, if they wish, suggest suitable referees for their articles.

TWO COPIES of scientific articles should be submitted. Authorities must be given to Latin names, either at their first mention or in a comprehensive list, and species lists should follow a named checklist. References should follow the so-called 'modified Harvard' convention, but with journal titles preferably given in full, not abbreviated. World List abbreviations will, however, be acceptable if the author indicates a definite preference. For guidance as to the preparation and presentation of material for publication, contributors are referred to the British Ecological Society's booklet A Guide to Contributors to the Journals of the BES, and The Royal Society's General Notes on the Preparation of Scientific Papers (3rd ed., 1974). Text proofs of papers will be provided and these should be returned to the Editors within three days of receipt.

Ten free reprints will normally be supplied to the authors of scientific articles; further copies will be available, if requested at the time proofs are returned, at a charge of 5p per side, plus postage.

Please note: there are no fixed deadlines for receipt of copy; material will normally be dealt with when received and will, if suitable, be published as soon as possible.

The Editors, Circsea, c/o Environmental Archaeology Unit, Walled Garden, University of York, Heslington, York YO1 5DD, U.K., Tel. 433848 (Harry Kenward), 433849 (Allan Hall) or Fax 433850.