1. The level of anonymisation that is necessary will depend on the research. The important point is that if data are not fully anonymised, participants should be aware of this and give their consent.

2. It is not required that access to the data be limited to the research team. Again, the key point is that if it is planned to make the data more widely available, then the participant should be told of this at the outset. Given that many research funders are encouraging data-archiving, it is a good idea to consider this at an early stage. The UK Data Archive notes that "The UKDA is concerned about the number of research projects that maintain they are unable to share research data due to promises made to research participants and interviewees to keep all research data confidential". For more information, see the website:

   http://www.data-archive.ac.uk/sharing/consentrecs.asp

3. When planning anonymisation measures, consider in what ways identification of participants may occur and what potential consequences identification would have. The aim is to maximize the security while minimizing the information loss. It is important not to over anonymise data. When research participants have been told upfront that the data would be archived for scientific reuse, data need to be anonymised to a level that ensures that re-users of data cannot immediately identify individual participants.

4. Research data may be used or disseminated for research purposes only. Handing over data to third parties or talking about individual participants to outsiders in a way that would affect the evaluation, treatment, status or behaviour of the participant is unethical.

5. Audio-visual datasets cannot be anonymised and so can be archived only if explicit consent was given for this.

6. In most cases the participants’ exact names and addresses should be destroyed after the original research has been completed. However, it may be necessary to retain contact information for longitudinal studies; again, the key point is to ensure that the participant has given explicit consent for this.

7. Where possible, other direct identifiers such as postcodes, telephone numbers, and exact birth dates should be removed from the data after the original research has been completed. Preserving them is justified only when direct identifiers are essential for the analysis of the data, and the participants have given specific consent to the arrangement beforehand.

8. Indirect identifiers/background information: The following are examples of background variables or indirect identifiers: gender, age, education, occupation, economic activity, socio-economic status, household composition, income, marital status, mother tongue, nationality, ethnicity, workplace/organisation, educational institution, and geographical identifiers. Geographical identifiers include, for instance, postcode, suburb, municipality, province, region, and place where the respondent grew up. The level of anonymisation needed depends on
whether a combination of indirect identifiers could lead to the identification of a respondent. If so, then variables can be recoded or deleted to avoid identification: for instance, instead of date of birth, age in years could be used; instead of a full postcode, use just the first 3 digits.

9. Responses to open-ended questions sometimes contain identifiers which are connected to respondents themselves or other persons, such as name or occupation of a spouse. Disclosure risk must be assessed on a case-to-case basis, with recoding, pseudonyms or deletion of variables being used if necessary to preserve confidentiality.

10. Less well-known anonymisation techniques include swapping and adding random variation to indirect identifiers. Swapping means matching unique cases on the indirect identifier and then exchanging the values of the variable. See:

   http://www.icpsr.umich.edu/icpsrweb/ICPSR/access/dataprep.pdf

11. Changing proper names to pseudonyms is the most popular anonymisation technique used for qualitative data. A good way to keep the anonymisation process under control is to replace personal names with pseudonyms directly after the transcription. Typing a special character in front of all proper names at the initial transcription stage will facilitate the planning and carrying out of anonymisation because all proper names can be easily found within the data.

12. A diagnosed severe illness can be changed into another, similar type of illness, if doing this does not reduce the usefulness of the data too much. Another option would be to categorise the information in the same way as with quantitative data. For example, 'AIDS' could be changed to [severe long-term illness] and thereafter referred to as [illness], provided that the reader is able to deduce from the context that [illness] refers to the 'severe long-term illness' mentioned at the beginning.

13. Detailed logs should be kept of all anonymisation measures carried out.

14. Besides data anonymisation, good research ethics can be enhanced by storing the data in a secure manner and setting conditions for its reuse.