Biomedical data: Ethical, legal and social implications

Sarah Morgan [1]

- Cross-domain
- Beginner
- 1 hour

This course explores the ethical, legal and social implications (ELSI) of biomedical data sharing. The material will introduce the reasons why we want to share data, the benefits it can bring, and the challenges of doing so. It will present ideas for best practice and cover resources that provide guidance and advice for sharing data.

Learning objectives:

- Introduce the basics of ELSI concerning biomedical data
- Awareness of ELSI resources

Biomedical data collection, use and sharing

Data, data everywhere...

Throughout our daily lives, a wealth of data is being collected about us - but much of this happens without us realising it. At the same time, many of us are happy to share data about ourselves through social media and other avenues. Much of the data being shared is personal in nature, yet for many people it isn't considered sensitive. However, when considering data that relates to biomedical study, linking through to an individual's health and wellbeing, many people consider this to be highly sensitive in nature.

Is biomedical data different?

We can therefore consider biomedical research data as a subset of personal, sensitive data and it must be treated in a particular manner. Nevertheless, it is important to be able to use, interrogate and share such data if it is going to be translated into information that may provide important insight into disease susceptibility, disease status, or response to specific medicines. Such information may benefit not only the individual who 'owns' the data, but also to the wider scientific community and the population at large.

Why is data sharing important?

The importance of sharing data is evident through the funding of projects such as BioMedBridges [2] and CORBEL [3].

The objective of these projects is to better enable researchers to access biomedical data and information by facilitating the process of sharing across organisations, scientific sectors and countries.

In their Ethical Governance Framework [4] document, the BioMedBridges project explains the importance of enabling data access, highlighting that increasing the use of data ultimately provides more societal benefit. The
example they provide is that of new health research discoveries "through the re-analysis of expensive, rare or unrepeatable investigations".

**What is ELSI?**

ELSI is defined as **E**thical, **L**egal and **S**ocial **I**ssues, explained below:

- **Ethical** - Is it ethical to collect the data? Is it being collected in an appropriate manner and used ethically?
- **Legal** - Is the data being collected, processed, stored and distributed in accordance with the legal requirements in place within that country or region?
- **Social** - Is data sharing important? What are the benefits and drawbacks of sharing this data? Is there a conflict between the rights of the individual and society as a whole?

Investigating and discussing the ethical, legal and social issues of biomedical data collection and sharing is an important activity as part of a research project. It provides a way of assessing the impact of the research activity on both the individuals from whom data is being collected and the society in which those individuals live.

The impact of data collection and sharing can be viewed both positively and negatively. Reviewing potential harms allows them to be avoided or minimised, whilst ensuring that positive outcomes are supported as fully as possible.

**Navigating the guidelines and regulations**

It can be extremely complicated to navigate the relevant ethical and legal requirements for collecting and using biomedical data, especially when you are looking to move data across national boundaries. Yet these are issues which researchers can face on a daily basis.

The aim of this module is to familiarise you with the ELSI considerations that need to be explored when planning and conducting research involving biomedical data; it will also link to resources developed by the BioMedBridges project that provide guidance on working with data and where relevant advice can be found.

**Where do issues arise?**

This module will consider three aspects of biomedical research data:

1. Collection.
2. Use (including storage and processing).

**Data collection**

Data is normally collected as part of a defined research project, in which a limited number of questions are to be addressed through the collection of a specific set of samples and data (or just data alone). Data can also be generated as part of the ongoing activity of a [biobank](5), where a collection of human-derived samples and data may be used to address any number of questions across a range of projects. Additionally, data can be collected through health services or government departments for use in healthcare planning.
Where data is to be collected for research purposes, in most cases ethical approval is required to undertake the research itself and (in general) participants must give their consent to be part of the study (or studies) in question.

Ethical principles

There are a great number of ethical/philosophical theories that can be applied when reviewing and discussing the ethics of biomedical research, but they can be grouped into three main types:

1. **Consequentialist theories** - the focus is on the consequence or outcome of an action. For example, if an action produces a good outcome, it is morally right.
2. **Deontological theories** - these identify moral duties that should be followed; the focus here is on the action that is taken, rather than the outcome.
3. **Virtue theory** - it is the moral character of the individual who is performing the action that is the focus; if an individual is acting virtuously then the actions of that individual (and hence the outcomes) will be morally right.

An approach that is widely applied in the field of biomedical ethics, however, is that of **principlism**, as described by Beauchamp and Childress in 1977. Principlism takes elements from a number of the theories described above, but provides four main principles that should be considered when undertaking ethical review:

**Respect for autonomy**: respecting the decision-making capacities of autonomous persons; enabling individuals to make reasoned informed choices.

**Beneficence**: actions must provide a form of benefit, whether a directly beneficial outcome, or producing the best overall result by balancing risks, benefits and costs.

**Justice**: there should be fair balance and distribution of risks and rewards.

**Non-malificence**: an obligation to not cause harm; whether physical (including the potential for death) or mental in nature.

Ethical review and approval

In order to conduct a biomedical research study involving human subjects and their data, approval is required from an ethics committee (known in the United States as an Institutional Review Board). This committee may act at an institutional level (for example a University Committee), or at a national level (such as The National Research Ethics Service in the UK).

Regardless of level, the role of these committees is the same: to review proposals for research projects and, through discussion amongst the group as a whole, determine whether these projects conform to the ethical norms within that society and do not contravene any legal requirements. Ethical norms vary across cultures, and what is deemed acceptable in one country may not be in another – even if they are close neighbours.

What is an ethics committee?

Within the European Union, an ethics committee can be defined as:

“an independent body in a member state of the European Union, consisting of healthcare professionals and non-medical members, whose responsibility is to protect the rights, safety and well being of human subjects...”

(Definition from the EU Clinical trials directive - Directive 2001/20/EC)
**Ethical review**

An ethics committee will review all aspects of a research study, including the full study protocol, participant information leaflets and consent forms, how data will be stored and used and how results will be disseminated post study.

For collection of new primary data, whether as a component of a project or the sole purpose, ethical approval will generally be required and informed consent will be sought from individuals to collect their data.

In the case of data that has been collected as part of a previous study, ethical approval may not be required. However, if the data is to be used for a different purpose from that originally agreed, it may be necessary to obtain additional informed consent from the individuals who provided the data in the original study. This additional consent may require new ethical approval.

**Informed consent**

Informed consent

For an individual to participate in a research project, they must provide informed consent. This is a principle that is central to the ethical undertaking of all biomedical research, and one that is enshrined in global human research guidelines and regulations via the Nuremberg Code (1947) and Declaration of Helsinki (World Medical Association, 1964; most recent update, 2013). Informed consent as a process is one in which both research participant and researcher (or the person taking consent if different from the researcher) need to take an active role. It can be broken down into 5 main steps:

1. Competency
2. Consent
3. Voluntariness
4. Understanding
5. Disclosure

**Competency**

Competency relates to an individual's ability to provide informed consent (competency is a legal definition) and if individuals are deemed to be non-competent then they will be unable to provide consent. They can, however, have the power of consent delegated to another individual such as a family member or a carer.

**Who is competent?**

Adults (often those over 18 years of age, but definitions may vary by country) who have no conditions which affect their ability to make sound decisions. The presence of mental health conditions and progressive disorders such as Alzheimer's disease may render participants non-competent.

Children are often described as non-competent (due to age) but in certain cases can still be allowed to take such decisions. For example, in the UK, children can be classed as Gillick competent, which allows them to give consent without parental/guardian interference.
Disclosure

To enable individuals to give fully informed consent, they must be fully apprised of all aspects of the study. This is the element of disclosure. All elements of the study process, including all potential risks and benefits, must be disclosed at this point, and the individual given the opportunity to ask questions.

How to inform?

Information is normally provided in the form of a written information sheet along with a verbal explanation from a member of the study team.

Alternative methods have been successfully used including audio recordings and filmed study overviews for people unable to read such literature.

Understanding

Providing participants with all the information is only one aspect of the informing process; the person taking consent must also assess whether the individual in question understands what they are signing up to. This ensures they have read the information, and know what their role in the study will be and what it entails.

Assessing understanding

Some projects have taken the approach of giving people an assessment before asking for consent (Personal Genome Project, Harvard) but in most cases this is done through a pre-consent conversation during which potential research subjects are asked various questions about the study.

Voluntariness

Individuals must provide consent by their own free will, free from coercion and any other influences. It is up to the individual taking consent to determine that the persons providing consent are doing so on their own behalf with no undue influence from other parties or external factors.

Coercive factors

Coercive factors can vary from culture to culture, although the two most commonly discussed are monetary influence and the potential for access to better healthcare provision if a clinical element is involved.

Individuals may also be influenced by family members or other key influencers, including medical staff.

Ethics committees look for potential coercive factors when reviewing a study for approval and may ask for these factors to be removed.

Consent

The provision of consent is usually the final step in this process. This is normally provided in a written form alongside the signature of the researcher (or other individual who is taking consent), although other methods can be used in cases where individuals are unable to write (for example, use of thumbprints).

Consent is, however, not always the final step, for two main reasons:
1. Should new information regarding the study become available, this will need to be communicated to the study participants and may require re-consenting if the changes are deemed to be significant.

2. Study participants may withdraw from a study at any point, in which case their consent is no longer valid.

**Study withdrawal**

Individuals who provide consent to be part of a research project have the right to withdraw from the study at any time if they no longer wish to be involved.

It can, however, be difficult to remove data already collected from a subject within a study. Therefore, individuals must be made aware that if they withdraw their consent, it may not be possible to remove this data.

**Data Use**

The use of personal data is controlled in many different ways and, as explained at the beginning of this course, finding out which rules apply to your research can be complex. In the European context, there are several layers of regulation and/or guidance that can be applied:

- **European level**: EU directive or direct EU legislation.
- **National (Country) level**: Local implementation of EU directive and country specific guidelines.
- **Organisation level**: Guidance for use of data within specific organisational settings.

The landscape of data regulation within Europe is potentially set to change with the introduction of direct European data legislation. However, as this is not currently enacted, the information provided within this course and in the BioMedBridges tools introduced under the ELSI resources section focus on current regulatory frameworks.

**EU directives**

In Europe, the protection of personal data is enshrined in both the European Convention on Human Rights (where Article 8 states that "everyone has the right to respect for his private and family life...") and more specifically in the Charter of Fundamental Rights in the European Union (2000/C 364/1) [6], where Article 8 focuses on the protection of personal data, with 8.1 stating "Everyone has the right to protection of personal data concerning him or her".

In an attempt to harmonise data laws across Europe, and to allow data flow across member states, a Data Protection Directive (95/46/EC) [7] was published in October 1995. This sets out the expected standards for data control within Europe, and applies to data that are processed by automated means and data that are part of non-automated systems but are accesible using specific criteria.

EU directives are enacted into each member state's legislative framework in accordance with its own laws and conventions. Therefore some differences exist between countries.

**The directive**

The directive sets out two 'roles' - that of the data controller (the individual or body who decides on the reasons and methods for processing) and the data subject (the individual to whom the personal data relates).
Data controllers have an important role in ensuring that data is processed fairly, and in accordance with what the data subject has agreed to. Data subjects have a number of rights in relation to their own data and must provide consent to its processing.

'Processing' in this context can include (but is not limited to): collection of personal data, its recording, storage, disclosure, consultation and adaptation.

**Data protection principles**

The data controller must adhere to the following principles:

- Data must be processed fairly and lawfully and must be collected for explicit and legitimate purposes and used accordingly.
- Data must be relevant and not excessive in relation to the purpose for which they are processed.
- Data must be accurate and, where necessary, kept up to date.
- Data controllers are required to provide reasonable measures for data subjects to rectify, erase or block incorrect data about them.
- Data that identify individuals must not be kept longer than necessary.
- Member States must provide one or more supervisory authorities to monitor the application of the directive.
- In principle, all data controllers must notify supervisory authorities when they process data.

**National guidelines**

Whichever member state you are undertaking research in, you need to determine what the national guidelines are within that country. Examples include:

**United Kingdom**

Data protection is legislated under the [Data Protection Act, 1998](https://www.legislation.gov.uk/ukpga/1998/40). The principles of data processing broadly follow the principles as set out within the 1995 Directive. Exemptions exist for use of certain data types for certain reasons, for example the use of health data (which is considered to be sensitive personal data) for public health reasons.

**Germany**

The [Bundesdatenschutzgesetz](https://www.bayern.de/daten-schutz-bayern) (Federal Data Protection Act) is the legislation that implements the directive under German law. There are, however, further laws, with each German state having its own data protection law.

**Additional safeguards**

In both countries there are additional safeguards for different data types. For example, in the UK’s National Health Service (NHS), Caldicott Guardians ensure the protection of data relating to NHS users in a single hospital. This involves the monitoring of data use beyond the reason for which it was initially collected, including research purposes.

**Data anonymisation**

Anonymisation of data removes its personal nature under the directive, and therefore changes its status in
What is data anonymisation?

**Complete data anonymisation** requires the complete removal of all personal identifiers from the data record so that an individual cannot be identified.

*What are identifiers?*

Identifiers are details that can uniquely identify the data subject, whether alone or when combined. These may include name, date of birth, gender, address, medical service number, and so on.

**Pseudonymisation or coding** involves attaching alternative identifiers to the data record, replacing the ‘real’ identifiers. For example, by changing names or removing some key identifiers and replacing them with a code. A cipher (which allows the breaking of the code to link the pseudonymised data to the original record) will be available.

This cipher may or may not be held by the organisation that holds the pseudonymised data, and in some countries (e.g. the UK and Germany), if the coded/pseudonymised data is held by one organisation and the cipher by another, it is not deemed to be 'personal' in nature. However, if one organisation holds both then this is not considered to be the case.

**Anonymise or not?**

Whether you fully anonymise your data will depend on what you plan to do with that data. In many cases for biomedical research, especially when involving a patient population, maintaining links to medical records is key to achieving the study outcomes. What is key from the data protection point of view is understanding where your data sits within the protection requirements.

Data sharing

Data sharing is becoming more common in biomedical research, and this sharing has both ethical and societal implications.

It can be considered unethical to collect data and not use it to its full potential. There are societal benefits to be gained from the greater dissemination of research findings. Yet, data sharing can present a number of potential issues, especially when it involves moving data between groups, organisations, or countries.

Of specific importance are:

- maintaining data confidentiality;
- maintaining data integrity.

Data confidentiality and data integrity

The process of sharing data can be direct from one researcher to another, or via externally available databases. These may be completely open access [10] (such as the majority of those provided by EMBL-EBI), or may be accessible only to individuals who request access via Data Access Committee (DAC) or similar (for example EMBL-EBI's European Genome–Phenome archive [11] or the Sanger Institute's Decipher [12]).

**Maintaining confidentiality and integrity**
How much access to the data do you have the ability to give (based on your ethical approval and informed consent of your research participants)?

If complete anonymisation is possible, can you pseudonymise and maintain the link-key within your organisation only?

It's crucial to ensure that your research participants have given consent for sharing of non-anonymised data, and that anyone with whom you share the data knows their responsibilities in maintaining it.

Sharing within the EU is safeguarded by the EU Data Protection Directive. However, safe sharing outside of EU member states requires an understanding of how data is treated within those countries/territories.

Data sharing agreements

In projects involving large consortia, especially those with consortium members distributed across the globe, data sharing agreements can be enacted.

The benefits of a data sharing agreement

These provide a framework and clear guidance on how data is to be shared within the project and the key process involved in doing so.

An example of such an agreement can be seen within the BioMedBridges Ethical Governance Framework [13], which sets out the principles for data use across the research infrastructures involved in the project.

Agreements such as these provide guidance and best practice based on the laws and regulations in the countries/global territories represented within the project partnership.

In certain fields of scientific research, such as genomics, international efforts are being made to produce codes of conduct applicable to all researchers in the field. A major example being the "International code of conduct for data sharing in genomics" initially proposed by Knoppers et al in 2011 (Towards a data sharing Code of Conduct for international genomic research, Genome Medicine 2011, 3:46) and published in 2014 by the Regulatory and Ethics Working Group, Global Alliance for Genomics & Health (The HUGO [14] Journal 2014, 8:1).

Case study - Biobanking

The following case study highlights some of the issues which can arise when working with data in particular domains.

Biobanks are research projects, often at a national level, that collect and analyse both data and samples that are collected longitudinally over a period of years.

The information held in a biobank is potentially very varied, and can include demographic details, health records, tissue and fluid samples and results of biomedical laboratory analyses.

Issues that may arise:

- Consent is often generic – biobanks are set up so that a range of investigators can use them as a resource to investigate a variety of diseases/processes etc.
- It is not always possible to state exactly what will happen to the information gained, therefore individuals must give broad consent to the use of such information.
- The samples and data held in such banks may need to be transferred across national boundaries to extend...
and enhance research studies.

A careful balance must exist between maintaining an individual's autonomy and stifling the potential research that could take place as part of the study.

Biobanking has been the focus of much discussion from an ELSI point of view, and whilst careful consideration is required when setting up such a project, a number of large national projects are successfully running across the EU. Research infrastructure projects such as BBMRI-ERIC [5] provide support for these national projects, facilitating access to the materials held within each country whilst also providing guidance on sharing materials and data without compromising the interests of the participants who provide the material.

BioMedBridges Ethical governance framework

One objective of the BioMedBridges project is to better enable researchers to access data, increasing its utility with the ultimate goal of benefiting society. Examples include facilitating new discoveries in health research and allowing re-analysis of expensive, rare or unrepeatable investigations. At the same time, BioMedBridges aims to continue to protect the interests of research participants with regards to their privacy and confidentiality.

The BioMedBridges Ethical Governance Framework [15] sets out policies for the project that specifically relate to ethical and regulatory issues surrounding the access of data used within the project. The aim of the framework is to enable the BioMedBridges project to operate within agreed terms with respect to participant consent, ethics committee approvals and national regulations, ensuring that researchers supply and access data based on a common ethical framework focussing on specific issues that are key to the development and operation of BioMedBridges.

As the project evolves, adjustments may be made to this framework. Any adjustments will be developed and agreed by the Ethical Governance Committee and approved by the Executive Steering Committee. Together with the Independent External Ethics Adviser, and in the context of deliverable 5.2 "Tool for assessment of ethical and regulatory requirements", project partners have also developed suitable participant information sheets and consent templates.

BioMedBridges is also working to improve the interoperability [16] between European research infrastructures. This work involves laying the groundwork that will facilitate the exchange, use and analysis of publicly-funded biomedical research data. Enabling researchers to make better use of existing experimental data is a sensible economic and ethical approach, allowing data re-use in different or new contexts. However, there are many legal hurdles to overcome because the data is biomedical in nature.

ELSI resources

BioMedBridges - Legal Assessment Tool

As part of the BioMedBridge project a Legal Assessment Tool [17] (LAT) was developed to help raise researcher awareness of the formal requirements for data sharing. The tool highlights areas that need further action from the researcher to make their data available, and issues prompts if expert advice is needed. A researcher can use the tool to learn more about requirements that cover the current legal framework for the European Union in the areas of data protection, data security, intellectual property and biosample security. It does not provide legally binding advice.
A researcher is prompted to answer a series of multiple choice questions, including questions about the type of data they want to share (metadata [18], text data, images, genetic data, biosamples or biosample associated data), the form in which the data is provided (data from which individuals can be identified or pseudonymised/anonymised data), or possible limitations on the wider use of the data (e.g. intellectual property requirements). After providing the necessary information, the user is shown the applicable rules and regulations for his or her specific case and the tool recommends possible solutions or necessary further steps to make the data shareable.

You can learn more about the LAT in the User Guide and Tool Description [19] document.
IPAC, the *International Policy interoperability and data Access Clearinghouse* [20], is part of the Public Population Project in Genomics and Society (P³G) consortium, based in Montreal, Canada. IPAC provides a ‘one stop’ screening service for policy interoperability and access authorisation. These non-legal services advise on research policies and protocols for data access; deliver a database of generic clauses and agreements that can be used to build ELSI policy documents; offer assistance in developing documentation, consent forms, code of ethics and policies for data sharing; provide a review service for proposed policies and procedure documents; present services for creating Data Access Compliance Office (DACO) for customised projects.

![IPAC Homepage](image_url)

**Figure 2** IPAC homepage

**REMS**

REMS, the *Resource Entitlement Management System* [21], provides a tool for the administration of access rights to research data. It acts as a broker between research scientists and the holders of biomedical data. Users can apply for access to datasets of interest through an application interface. REMS then notifies the data custodian/s, as well as presenting the researcher with the documentation required for access.

Further information on REMS can be found on the REMS website. This includes leaflets, presentations, literature, a demonstration of the tool and information on the REMS software.

**BBMRI Legal Wiki**

The European *Biobanking and Biomolecular Resources Research Infrastructure* [5] (BBMRI), and BioMedBridges partner, has created a wiki to provide a community resource for disseminating legal materials to the BBMRI community. It includes legal forms and standards, information on navigating legal pathways, and guidance on using legal templates. The *BBMRI Legal Wiki* [22] is intended to foster discussion and the sharing of experiences by BBMRI members and partners, making it easier to access experience and expertise on a complicated subject.
Your feedback

Please tell us what you thought about this course. Your feedback is invaluable and helps us to improve our courses and thus enhance your learning experience.

References

- Council directive 2001/20/EC on the approximation of the laws, regulations and administrative provisions of the Member States relating to the implementation of good clinical practice in the conduct of clinical trials on medicinal products for human use [2001].
- WMA Declaration of Helsinki - Ethical Principles for Medical Research Involving Human Subjects, 7th Revision, 2013.
- European convention on human rights [1950], Article 8.
- Council directive 95/46/EC on the protection of individuals with regard to the processing of personal data and on the free movement of such data [2005].

Contributors

Sarah Morgan [1]
EMBL-EBI
Training Coordinator

Sarah has a BSc and MSc in Biomedical Sciences, a PhD from Cranfield University and an MA in Healthcare Law and Ethics from the University of Manchester. Over a 10 year academic career her research has focused on tumour biomarker characterisation and cell-surface interaction. Beyond research she took an active role in the provision of postgraduate taught courses for the University as Lecturer in Molecular Medicine, leading the development and direction of a number of MSc Courses including the MSc in Translational Medicine, the MSc in Molecular Medicine, and more recently acting as director for the MSc Programmes in Advanced Biosciences. She has lectured on a variety of biomedical techniques and topics, including ethics and governance of research.

Tom Hancocks [25]
EMBL-EBI
Scientific Training Officer - BioMedBridges

Tom Hancocks works as a Scientific Training Officer for the Training Team at EMBL-EBI.

He studied Human Genetics at the University of Leeds and McMaster University in Hamilton, Ontario; before completing an MSc in Analytical Genomics at the University of Birmingham.

Tom has worked for the NHS in diagnostic genetics and as a bioinformatics trainer for healthcare scientists and clinicians.

Funding

This e-learning module was developed as part of the BioMedBridges [2] project, funded by FP7 Capacities Specific Programme, grant agreement no. 284209.
Biomedical data: Ethical, legal and social implications
Published on EMBL-EBI Train online (https://www.ebi.ac.uk/training/online)

Links
[1] https://www.ebi.ac.uk/training/online/trainers/sarahm
[10] https://www.ebi.ac.uk/training/online/glossary/open-access
[11] https://www.ebi.ac.uk/training/online/course/ega-quick-tour
[12] https://decipher.sanger.ac.uk/
[13] https://www.ebi.ac.uk/training/online/course/elsi/biomedbridges-ethical-governance-framework
[14] https://www.ebi.ac.uk/training/online/glossary/hugo
[16] https://www.ebi.ac.uk/training/online/glossary/interoperability
[18] https://www.ebi.ac.uk/training/online/glossary/metadata
[21] https://confluence.csc.fi/display/REMS/Home
[23] http://dx.doi.org/10.1089/bio.2015.0014
[24] https://www.ebi.ac.uk/training/online/glossary/biomarker
[25] https://www.ebi.ac.uk/training/online/trainers/Tom