

Wellcome Genome Campus | 3-9_Data sharing best practices

When dealing with data, we need to remember some of the best data-sharing practises. It is agreeable that sharing of data and, besides, sources is essential for advancing health, expedites translation of research results into knowledge, products, and procedures to improve human health. Also, timely release and sharing of data, such as clinical data, genomic, and any other data types associated with the research samples is very key.

All the data sets generated ought to be made freely available via deposition. This can be into publicly accessible and searchable international databases which include but not limited to the GenBank; the National Centre for Biotechnology Information, also known as NCBI; the European Genome-Phenome Archive, commonly known as EGA; or any other public databases. Key patient information, therefore, should be maintained with utmost confidentiality. That data should be stripped of identifiers prior to release for sharing.

Also, the data and its associated documentation should be made available to only the users who agree to the terms and conditions of a data access and sharing agreement. This is very important, as it aims to protect the privacy and interests of the research participants.

The agreement must, therefore, require to agree to, one, use of the data only for the approved research, to protect the participants' confidentiality, to follow appropriate data security protections. They ought to follow applicable laws, regulations, and local institutional policies and procedures for handling of the data, not to attempt to identify individual participants from whom the data within a data set were obtained.

Confidentiality is, therefore, very key. They are not to sell any of the data elements from the data sets obtained, not to share individuals other than those listed in the request, and then to point out any problems or concerns about the data. Lastly, destroying or returning the data after analyses are completed. Thank you.

Also important to note, anyone intending to share research data must develop guidelines that are consistent with their local regulatory authority guidelines, those of the funders, and that make use of the NIH and NIAID data-sharing guidelines, as well as being consistent with the contemporary principles such as FAIR, which is Findable, Accessible, Interoperable, and Reproducible standards for data release. The principles for data sharing and access may include maximising the availability of research data in a timely and responsible manner, protecting the rights and privacy of human subjects who participated in the research studies, recognising the scientific contribution of researchers who generated the data, and considering the nature and risk aspects of the proposed research whilst ensuring the timely release of this data, and promoting deposition of data in the existing community data repositories whenever possible. Thank you.