

Submission of Evidence to the Genomic Medicine House of Lords Committee from the European Bioinformatics Institute

Summary

The use of molecular assays to inform clinical decisions shows great promise and will revolutionise medicine over the coming decades. In particular determining an individual's genomic signature, at first by genotyping likely to be replaced by genome sequencing will provide the key dataset to inform on the genetic contribution to common disease risk, specific high penetrant diseases and response to treatment. Other molecular assays include the use of microarrays, for example in tumour samples, or blood biomarkers for diagnostic use.

In all these cases, the assay technology will generate large amounts of data specific to that individual. Each data type has its own unique challenge in storage, manipulation and interpretation, of which there is nearly always an active research community who have used this assay for basic and applied research. It is therefore critical that the knowledgebase created for basic research continues to develop and that this knowledgebase is leveraged appropriately for the use of molecular assays in the clinic. The presence of large scale molecular data in the clinical setting will require the establishment of sophisticated information management groups, involving new management structures and new recruitment.

An important difference between basic research and clinical information is the private nature of clinical information to the patient and trusted healthcare professionals. This implies that information which is only used for an individual's medical care should remain private. However, the storage, manipulation and interpretation of this data should be coordinated with the basic research systems. In addition, there are datasets of clinical information which have been deliberately collected to advance clinical research, such as those underlying genetic disease association or epidemiological studies, such as UK BioBank. In this case, a form of restricted access, consistent with well established medical ethics, has recently been developed.

In the setting of basic research best practice has been to build large, open datasets of the molecular assay results, usually coordinated worldwide. These datasets are then integrated to provide a richer description of the biological processes involved. In Europe, the European Bioinformatics Institute, which is sited in south Cambridgeshire, Hinxton, is the recognised centre for this activity and holds all of the major biomolecular data resources. The EBI has the potential to coordinate activity inside the UK, among European member states and internationally. The UK government should take advantage of these European structures, in particular given the physical location

In more detail, we recommend a three tier system, outlined below.

Support of Basic Research base

The use of molecular data in the clinical setting relies on a vibrant and open fundamental research base. In the area of bioinformatics, many of the proposed molecular data types are being handled, with data storage and archiving occurring primarily at the EBI. It is critical therefore that the basic research continues to be supported. However, the growth in molecular assay generation has often outstripped the provision of funds for the aggregation and utilisation of this data, in particular because many funding instruments are not ideally designed to work with such infrastructure components. Recognising this, the European Strategy Forum on Research Infrastructures (ESFRI) launched an in depth, bottom up preliminary phase for scientific infrastructures. One of the 36 infrastructures currently being considered is **ELIXIR**, the European Lifesciences Infrastructure for Biological Information, lead by Professor Thornton. The UK government has an opportunity to show crucial leadership to support this infrastructure which underpins these clinical applications, and thus also leverage European funds for this endeavour.

Handling personal genomic and molecular assay records

Unlike basic research, information on individual patients cannot be openly released. This information conceptually is part of the medical record of the patient, but its form and interpretation are far more complex than text based records of diagnostic decisions. It is important that the NHS IT system considers the management of these complex datasets as part of their roadmap, and that these systems are being built with appropriate input from the basic research domain to maximise the utility of the data and minimise the risk in the development of new systems. This is likely to involve transfer of both software systems and personnel, as well as training in informatics in the clinical setting.

Support of the clinical knowledgebase

The third main information component will be the knowledgebase of clinically relevant interpretations of the biomolecular assays. Conceptually part of this is akin to the verification of treatment regimes and new pharmaceutical compounds by institutions such as NICE. This area is not completely understood, and there is considerable clinical research required to define the best application of this molecular data. An additional requirement is that the output of this research can be combined with the personal medical record, which would include the molecular data, to provide concise, evidence based recommendations for medical decisions. This latter aspect has a complex informatics component, somewhat akin to the concept of “medical algorithms” but applied to molecular data. In addition, for medical professionals to maximise this information there needs to be both specific training in tools and a broad understanding of the concepts and assays in the medical training curriculum.