Establish good genomic practice to guide medicine forward

Richard W Barker[1], David A Brindley[1–3] & Anna Schuh[4]

Genomic advances, including next-generation sequencing, offer substantial opportunities and challenges for stratified and personalized medicines. However, the lack of standardization in genomic diagnostics translates into a major risk of error introduction. To ensure the integrity of such data—and their application—we suggest the development of 'good genomic practice' standards to guide the field.

Genomic analysis, which plays an increasing part in patient stratification and prediction of treatment outcomes, follows a multistep and multidisciplinary process spanning the initial sample collection to the ultimate clinical decision. However, a risk of error introduction exists at all stages, from preanalytical steps to laboratory analysis to clinical interpretation and long-term data storage. At present, good laboratory practice (GLP) and good manufacturing practice (GMP) guidelines ensure the quality of tests and therapeutic agents, thus bolstering the reproducibility and consistency of medical research and practice. But the emergence of clinical genomics poses new and specific challenges, and, as such, we believe that good genomic practice (GGP) standards would represent a similarly vital step in the field's development.

Before a genomic analysis even begins, the clinical and ethical implications of genomic testing, such as the possibility of incidental findings revealing the presence of heritable diseases, should be assessed as part of GGP (*Nature* **429**, 478–481, 2004). Clinical genomics training must be considered, as well as, where necessary, clinical service reconfigurations—for example, between genetic counseling and disease-specific clinical specialties (*Nature* **470,** 204–213, 2011).

Maximizing tissue sample quality and minimizing intrasample heterogeneity is crucial. For example, biopsies containing mixtures of normal and tumor cells require review and macrodissection by skilled staff. These considerations are all the more important today given the progressive size reductions in biopsies taken from patients. At the same time, for diseases such as cancer, longitudinal studies and sampling of metastatic material may need to become routine. Given the importance of such sampling, a reliable chain of custody from tissue acquisition to initial analytical processes is crucial. For example, centralized specimen handling accommodating paired tumor and germline samples, barcoding and radio-frequency identification tracking systems to ensure sample security, traceability and information technology connectivity from the patient to the final diagnostic report are all essential.

At the analytical stage, DNA and RNA library preparation methods should be optimized as well as standardized to accommodate the quality of input material and ultimate results. Standard operating procedures should define correct positive and negative controls. These procedures should also minimize contamination risks and interrun and interlaboratory variability. Quality measures must define process failures and duplicate testing. Additionally, labs should use fully validated GMP-grade reagents when available.

The wide range in bioinformatics tools is a major source of the discrepancies in genomics results when it comes to determining the frequency and significance of a particular sequence variation. Validation and standardization of version-controlled alignment tools, variant calling and tumor-normal subtraction algorithms will be required to achieve consistency in reporting relevant mutations.

It is necessary to rigorously curate publicly available databases of sequence variations to determine the presence of clinically relevant variants. Importantly, the final step of any bioinformatics pipeline that aims to produce 'medical-grade genomes' should be the visual inspection of sequencing data by skilled staff.

Assessing the clinical significance of variants remains a key component, including how to manage their medical impact. The level of evidence regarding a variant's clinical significance can range from prospective randomized, controlled multicenter trials to retrospective analyses of single-center cohorts to pedigree studies to functional *in vitro* data. And the clinical bearings of variants can be diverse: they can change a diagnosis, predict prognosis without changing clinical management or serve as strong predictors of treatment response. Thus, stakeholders should work toward consensus guidelines on how to categorize pathogenic variants according to their clinical significance.

Over time, the scope of best practices, including GLP and GMP, has extended beyond the doors of processing suites, for example including sample shipping conditions or posting of clinical reports. Similarly, GGP will have to extend to long-term storage and use of patient genomic data in phenotype and genotype databases. And it should stipulate secure data storage standards, encryption and anonymization to preserve patient confidentiality, where appropriate. GGP should also define the type of genomic information that has to be stored and for how long. Additionally, it should encourage the development of internationally accepted and standardized data storage formats, enabling the secure transfer of data from different laboratories and thus reducing the likelihood of research duplication and supporting validation.

To implement the GGP concept, we suggest the creation of a checklist by an international genomics leadership group that highlights any extant standards that should be used and any gaps in the process that require a concerted standardization effort. Thereafter, pursuing further development through engagement with a wider stakeholder group, a GGP accreditation pathway should be proposed. This accreditation might be used, for example, to help ensure the reproducibility of exome sequencing studies and to support regulatory interactions and submissions.

Genomic medicine is a young discipline with boundless potential to improve human health. We strongly believe that it is timely and crucial for the genomics community to define essential attributes of GGP and the 'medical-grade genome', thus ensuring the field realizes this great potential.

[1]Oxford–UCL Centre for the Advancement of Sustainable Medical Innovation, University of Oxford, Oxford, UK. [2]Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences, University of Oxford, Oxford, UK. [3]The Harvard Stem Cell Institute, Cambridge, Massachusetts, USA. [4]Oxford Biomedical Research Centre Haemato-Molecular Diagnostic Centre, Oxford, UK.