

ANALYSIS

Policy challenges of clinical genome sequencing

Around the world, genome sequencing is moving from research into the clinic, and in the UK plans to sequence the genomes of 100 000 NHS patients are well underway. A clear policy on how to conduct genomic testing is therefore both essential and urgent, argue **Caroline Wright and colleagues**

Caroline F Wright *senior scientific manager*¹, Anna Middleton *ethics researcher*¹, Hilary Burton *director and public health consultant*², Fiona Cunningham *Ensembl variation project leader*³, Steve E Humphries *professor of cardiovascular genetics*⁴, Jane Hurst *consultant clinical geneticist*⁵, Ewan Birney *associate director*³, Helen V Firth *consultant clinical geneticist*⁶

¹Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, UK; ²PHG Foundation, Cambridge, UK; ³European Molecular Biology Laboratory, European Bioinformatics Institute, Hinxton, UK; ⁴Centre for Cardiovascular Genetics, UCL Institute of Cardiovascular Science, University College London, London, UK; ⁵North East Thames Genetics Service, Great Ormond Street Hospital, London, UK; ⁶Cambridge University Hospitals Foundation Trust, Addenbrooke's Hospital, Cambridge, UK

Genetic testing is moving from analysis of specific genes to sequencing of the whole genome. Clinical genome-wide sequencing is already offered by a handful of private companies and diagnostic laboratories in the US and by some countries in Europe. In December 2012, the UK prime minister announced ambitious plans to sequence the whole genomes of 100 000 NHS patients over the next three to five years.¹ And in July the Department of Health set up Genomics England to help deliver the 100K Genome Project into mainstream healthcare in the NHS, with the initial focus on patients having genetic testing for the diagnosis of rare disorders, cancers, and infectious disease.²⁻⁶ Policy makers around the world are currently grappling with how to guide the implementation of genome sequencing in the clinic. Clear testing policy now needs to be agreed that covers issues such as whom to test and how to store, protect, and share genomic data appropriately.

Why sequence genomes rather than genes?

Traditionally, molecular genetic tests have involved sequencing single genes. Although this has proved to be a powerful method for diagnosing patients with rare heritable diseases, it has important limitations; a clinician has to select just one gene to test from the roughly 20 000 in the human genome. For a few rare disorders, such as cystic fibrosis, the clinical presentation is sufficiently distinct to make this possible. But many conditions are genetically heterogeneous, meaning that the same clinical picture can be caused by mutations in any one of many genes—for example, hypertrophic cardiomyopathy. Moreover, small but important subsets of common diseases can be caused

by a single gene defect (such as MODY in diabetes), but it may be impossible to identify these based solely on the clinical presentation. Testing individual genes sequentially is slow, expensive (costing hundreds of pounds each), labour intensive, and often ultimately unsuccessful. All this makes the one size fits all approach of testing every gene in the genome at a competitive price (box 1) an attractive diagnostic option, as well as facilitating research to find the cause of hitherto undiagnosed genetic conditions.

The diagnostic power of genome sequencing, coupled with our increasing understanding of the genetic aetiology of numerous disorders, seems to offer new opportunities to prevent, diagnose, and manage diseases. However, although scientific knowledge and technologies continue to advance rapidly, there are numerous outstanding questions surrounding clinical implementation. Whom should we sequence? How much of the genome should be sequenced? What should be tested, validated, and communicated to patients? How should individuals' genomic data be securely stored for their own benefit while being widely shared for everyone's benefit? Opinions differ dramatically,^{8,9} and a clear, evidence based genomic testing policy now needs to be agreed as a matter of urgency for the NHS.

Problem of data overload

The declining cost of DNA sequencing has shifted the diagnostic bottleneck away from performing the assay (that is, determining a genetic sequence) towards interpreting the test (predicting the effect of genomic variants in an individual). Our ability to

Box 1: What we know (and don't know) about the human genome

- Sequencing a human genome—the complete genetic information encoded by a person's DNA—currently costs around £4000 (€5000; \$6000)
- The exome is a subset of 1-2% of the genome that includes all protein coding genes (about 20 000); sequencing currently costs £200-500
- Over 80% of the genes in the exome currently have no known role in disease
- The non-coding sequence that comprises the remaining 98-99% of the genome is even harder to interpret
- Most people have around 3-4 million variants (sites where an individual's sequence differs from the reference genome), of which 30 000 to 60 000 are in the exome, including around 400 potentially damaging variants and, on average, two disease causing mutations⁷
- Most currently known clinically relevant variants are located in the exome

generate data now far outstrips our ability to interpret it. However, like with many other tests in medicine, the diagnostic accuracy varies according to the clinical history and pre-test probability.

One of the big issues we face is ambiguity around how to interpret genomic variants in an individual, particularly when he or she has no known family history or symptoms of a disease. When a patient presents with a particular phenotype, and a pathogenic mutation in a relevant gene is found, this can be identified as the cause of the disease with reasonable certainty. However, contrary to popular belief, the reverse is not always true. People can carry a disease causing variant without having the disease, either because of incomplete penetrance (not everyone carrying a particular genetic mutation will develop the disease) or age dependent penetrance (the disease does not manifest until later life). Without knowledge of the effect of mutations in the general population, we simply do not know how likely an unaffected individual carrying a particular mutation is to manifest that disease. Moreover, it is often difficult to predict the severity and course of disease in individuals even for well characterised variants in well known genes.

Figure 1⇓ relates to genome analysis of a patient, which confirmed the clinical suspicion that he had Charcot-Marie-Tooth disease.¹⁰ This diagnosis could have been made by analysing the sequencing results for a panel of just 40 genes known to cause neuropathy to identify the two relevant variants. These variants could be interpreted with confidence in the knowledge that the patient had a clinical diagnosis of an inherited neuropathy, making the odds that a variant in a neuropathy gene will be disease causing much higher than in a member of the general population.

But what about the rest of his genome? Analysis identified 159 known variants in other genes that are associated with various different diseases and traits—often termed incidental findings. What does this mean for him? This question is much more challenging to answer, and we must beware of overinterpreting the data. For example, this adult patient has a homozygous mutation in the *IGHMBP2* gene previously thought to cause a disease that is fatal in infancy (spinal muscular atrophy with respiratory distress type 1^{11 12}). Unfortunately, our understanding of this and many other genetic variants is often incorrect, and, had he been an infant when his genome was sequenced, exploring this result would have wasted valuable resources and caused enormous distress.

The problem with sequencing an individual's genome is that it will inevitably create many such difficulties in every single case unless a strategy is developed to partition the data and interrogate subsets of it according to the presenting clinical problem.

Incidental findings—screening by another name?

The question of how to handle incidental or secondary findings arising from genomics is a topic of heated international debate.¹³⁻¹⁶ Questions relate to whether, when, and what to report back to patients and research participants, how the process could or should be managed, and the inevitable resource implications. Although professional organisations have now made recommendations (table⇓),^{17 18} these differ on important areas such as screening in children and remain highly controversial.¹⁹

Sequencing a genome does not equate to screening a genome, and we can choose to limit our analysis to any number of specific variants or diseases. For clinical genome sequencing, we suggest that the detailed analysis should be limited to the pertinent genes—that is, those likely to be relevant to the disease phenotype under investigation (fig 2⇓). Analysing the non-pertinent genes can give rise to a potentially large number of opportunistic findings and can be considered similarly to genomic screening.²⁰ The crucial difference between pertinent diagnostic findings and opportunistic screening is the prior probability of disease in that individual—that is, the likelihood of the diagnosis before testing based on clinical findings—which affects how the results are interpreted.

When considering whether to offer any form of genomic screening, we must bear in mind the experience of existing screening programmes: it is easy to overstate potential benefits while underestimating potential harms, and well established and internationally recognised screening criteria should be applied when considering what (if any) additional genomic screening should be offered to those undergoing sequencing. Importantly, each variant-disease relation must be evaluated separately in light of the evidence associating them and the availability, invasiveness, and cost of confirmatory testing and treatment.

Well characterised and easily treatable diseases for which the benefits of cascade screening of relatives have been established could be initial candidates for opportunistic genomic screening—for example, DNA based screening of relatives of patients with familial hypercholesterolaemia has been shown to considerably improve cost effectiveness of diagnosis and health.²¹ Integration of genetic testing with established screening programmes should be formally evaluated before such data are used for this purpose. Screening for breast cancer susceptibility variants, for example, could be combined with existing risk data to target mammographic surveillance at women most at risk.^{22 23}

Widening public understanding

Professional and public opinion on whole genome sequencing varies widely, and the results of ongoing large scale social science studies exploring this area should ultimately contribute to policy.²⁴ Some argue that anyone undergoing genomic testing

should have the right to see all the test results, even when their clinical importance is uncertain. Although this position may seem to respect individual autonomy, it does not resolve the question of whether healthcare providers have a duty to search for and act on clinically actionable variants, or indeed to interpret data with uncertain predictive value that is not directly pertinent to the current clinical question (or the clinician's immediate expertise). And what of the autonomy of relatives, who may be affected by clinically actionable variants?

Individuals differ substantially in their tolerance for uncertainty, and the challenge of interpretation is underappreciated by both clinicians and the public, many of whom still regard genetic information as highly deterministic.

Questions also arise over how to balance the need for data sharing for the common good with the importance of respecting an individual's right to personal privacy and confidentiality. Clarity over legal and professional responsibilities are needed so that clinicians are not driven to overinvestigate because of fear of repercussions.

The challenge lies in ensuring that patients are aware of the issues surrounding genome sequencing, the consequences of results, and procedures for data sharing and storage before agreeing to participate. In the recently launched Personal Genome Project UK, the healthy volunteers undergoing genomic sequencing must pass an online examination testing their understanding of genetics and sign a lengthy consent form.²⁵ Exactly how the informed consent process will work for the government's 100K NHS genome project remains uncertain.

Focus on diagnosis

Given the limitations in our understanding of genomic variation, intelligent use of sequencing technology demands that we select which parts of the massive amounts of data we want to interrogate in detail to answer a clinical question. However, the remainder of the data should be made available for research. An informed, targeted approach to genome analysis makes the clinical test a more discrete and definable entity that is possible to interpret and reduces unwanted incidental findings. As knowledge grows, additional portions of the data could be interrogated as appropriate.

We therefore suggest that clinical genome sequencing efforts should initially focus on delivering diagnoses for patients rather than premature opportunistic screening. By combining genome wide sequencing with clinically targeted analysis (or gene panel tests) in patients whose clinical presentation suggests a primarily genetic aetiology, we can maximise interpretable, pertinent findings and minimise non-pertinent, incidental findings. Before genomic screening is considered, we need systematic, longitudinal investigation of variants in large populations to determine penetrance, and this could be done through ongoing biobank sequencing projects. Guidance and educational material for clinicians and patients involved in genomic sequencing in the NHS should be developed, covering the purpose and limitations of testing, the uncertainties of interpretation, and the importance of sharing data.

In addition, all variants ascertained clinically or through research should be deposited in accessible databases of genomic variation. To facilitate this, we need international collaborative efforts to systematically record and share genotypic and phenotypic information,²⁶ such as the DECIPHER consortium,²⁷ the Leiden Open Variation Database,²⁸ or the Locus Reference Genomic collaboration.²⁹ This will allow us to assemble sufficient evidence to reap the benefits of the genomic revolution responsibly and effectively.

We thank European Bioinformatics Institute for supporting the meeting which gave rise to this article, John Firth for comments on the manuscript, and all the workshop attendees and presenters for their thoughtful contributions, including Stephanie Barton, David FitzPatrick, Matthew Hurlles, Eamonn Maher, Helen Parkinson, Justin Paschall, and Vincent Plagnol. SEH is supported by the British Heart Foundation.

Contributors and sources: The authors were all participants at an informal meeting of multidisciplinary experts held in April 2013 at the Wellcome Trust Genome Campus in Cambridge, UK, to discuss the interpretation of human disease related variants. All are directly involved in genomic research, delivery of NHS genetic services to patients, or developing genomics health policy.

Conflicts of interest: We have read and understood the *BMJ* policy on declaration of interests and declare the following relevant interests: EB consults for Oxford Nanopore, a next generation sequencing company.

Provenance and peer review: Not commissioned; externally peer reviewed.

- 1 H M Government. Strategy for UK life sciences: one year on. 2012. www.bis.gov.uk/assets/biscore/innovation/docs/s/12-1346-strategy-for-uk-life-sciences-one-year-on.pdf.
- 2 Worthey EA, Mayer AN, Syverson GD, Helbling D, Bonacci BB, Decker D, et al. Making a definitive diagnosis: successful clinical application of whole exome sequencing in a child with intractable inflammatory bowel disease. *Gen Med* 2011;13:255-62.
- 3 De Ligt J, Willemsen MH, van Bon BWM, Kleefstra T, Yntema HG, Kroes T, et al. Diagnostic exome sequencing in persons with severe intellectual disability. *N Engl J Med* 2012;367:1921-9.
- 4 Leary RJ, Kinde I, Diehl F, Schmidt K, Clouser C, Duncan C, et al. Development of personalized tumor biomarkers using massively parallel sequencing. *Science Transl Med* 2010;2:20ra14.
- 5 Diamandis M, White NMA, Yousef GM. Personalized medicine: marking a new epoch in cancer patient management. *Mol Cancer Res* 2010;8:1175-87.
- 6 Köser CU, Holden MTG, Ellington MJ, Cartwright EJP, Brown NM, Ogilvy-Stuart AL, et al. Rapid whole-genome sequencing for investigation of a neonatal MRSA outbreak. *N Engl J Med* 2012;366:2267-75.
- 7 Xue Y, Chen Y, Ayub Q, Huang N, Ball EV, Mort M, et al. Deleterious- and disease-allele prevalence in healthy individuals: insights from current predictions, mutation databases, and population-scale resequencing. *Am J Hum Genet* 2012;91:1022-32.
- 8 Burn J. Should we sequence everyone's genome? Yes. *BMJ* 2013;346:f3133.
- 9 Flinter F. Should we sequence everyone's genome? No. *BMJ* 2013;346:f3132.
- 10 Lupski JR, Reid JG, Gonzaga-Jauregui C, Rio Deiros D, Chen DCY, Nazareth L, et al. Whole-genome sequencing in a patient with Charcot-Marie-Tooth neuropathy. *N Engl J Med* 2010;362:1181-91.
- 11 Tachi N, Kikuchi S, Kozuka N, Nogami A. A new mutation of IGHMBP2 gene in spinal muscular atrophy with respiratory distress type 1. *Pediatr Neurol* 2005;32:288-90.
- 12 Lupski J, Gonzaga-Jauregui C, Yang Y, Bainbridge M, Jhangiani S, Buhay C, et al. Exome sequencing resolves apparent incidental findings and reveals further complexity of SH3TC2 variant alleles causing Charcot-Marie-Tooth neuropathy. *Genome Med* 2013;5:57.
- 13 Wolf SM, Lawrenz FP, Nelson CA, Kahn JP, Cho MK, Clayton EW, et al. Managing incidental findings in human subjects research: analysis and recommendations. *J Law Med Ethics* 2008;36:219.
- 14 Wolf SM. The past, present, and future of the debate over return of research results and incidental findings. *Genet Med* 2012;14:355-7.
- 15 Bredenoord AL, Kroes HY, Cuppen E, Parker M, van Delden JJM. Disclosure of individual genetic data to research participants: the debate reconsidered. *Trends Genet* 2011;27:41-7.
- 16 Gliwa C, Berkman BE. Do researchers have an obligation to actively look for genetic incidental findings? *Am J Bioethics* 2013;13:32-42.
- 17 Green RC, Berg JS, Grody WW, Kalia SS, Korf BR, Martin CL, et al. ACMG recommendations for reporting of incidental findings in clinical exome and genome sequencing. 2013. www.acmg.net/docs/ACMG_Releases_Highly-Anticipated_Recommendations_on_Incidental_Findings_in_Clinical_Exome_and_Genome_Sequencing.pdf.
- 18 Van El CG, Cornel MC, Borry P, Hastings RJ, Fellmann F, Hodgson SV, et al. Whole-genome sequencing in health care. *Eur J Hum Genet* 2013;21:580-4.
- 19 Holtzman NA. ACMG recommendations on incidental findings are flawed scientifically and ethically. *Genet Med* 2013;15:750-1.
- 20 Wright CF, Burton H, Hall A, Moorith S, Pokorska-Bocci A, Sagoo GS, et al. Next steps in the sequence: the implications of whole genome sequencing for health in the UK. PHG Foundation. 2011.
- 21 Nherera L, Marks D, Minhas R, Thorogood M, Humphries SE. Probabilistic cost-effectiveness analysis of cascade screening for familial hypercholesterolemia using alternative diagnostic and identification strategies. *Heart* 2011;97:1175-81.
- 22 Evans DGR, Warwick J, Astley SM, Stavrinou P, Sahin S, Ingham S, et al. Assessing individual breast cancer risk within the UK national health service breast screening program: a new paradigm for cancer prevention. *Cancer Prev Res* 2012;5:943-51.
- 23 Burton H, Chowdhury S, Dent T, Hall A, Pashayan N, Pharoah P. Public health implications from COGS and potential for risk stratification and screening. *Nat Genet* 2013;45:349-51.
- 24 Middleton A, Parker M, Wright CF, Bragin E, Hurlles ME, DDD Study. Empirical research on the ethics of genomic research. *Am J Med Genet* 2013;161:2099-101.
- 25 Tørljensen I. Genomes of 100 000 people will be sequenced to create an open access research resource. *BMJ* 2013;347:f6690.
- 26 Creating a global alliance to enable responsible sharing of genomic and clinical data. 2013. www.sanger.ac.uk/about/press/assets/130605-white-paper.pdf.
- 27 Bragin E, Chatzimichali EA, Wright CF, Hurlles ME, Firth HV, Bevan AP, et al. DECIPHER: database for the interpretation of phenotype-linked plausibly pathogenic sequence and copy-number variation. *Nucleic Acids Res* 2013 Oct 22 [Epub ahead of print].

Key messages

Genome sequencing is now sufficiently cost effective to be offered clinically

Nevertheless, interpretation of individual genomic variation remains challenging and the importance of incidental findings is unclear

It is premature to offer opportunistic genome screening without knowledge of population penetrance

Clinical genome sequencing efforts should initially focus on delivering diagnoses for patients

A shared database of sequencing results linked to phenotypes is needed to facilitate research

- 28 Fokkema IFAC, Taschner PEM, Schaafsma GCP, Celli J, Laros JFJ, den Dunnen JT. LOVD v.2.0: the next generation in gene variant databases. *Hum Mutat* 2011;32:557-63.
- 29 Dagleish R, Flicek P, Cunningham F, Astashyn A, Tully R, Proctor G, et al. Locus reference genomic sequences: an improved basis for describing human DNA variants. *Genome Med* 2010;2:24.

Accepted: 8 October 2013

Cite this as: [BMJ 2013;347:f6845](#)

© BMJ Publishing Group Ltd 2013

Table

Table 1 | Key differences between the American College of Medical Genetics and Genomics (ACMG) and European Society of Human Genetics (ESHG) recommendations regarding whole genome sequencing in healthcare

ACMG ¹⁶	ESHG ¹⁷
Diagnostic laboratories should routinely screen all clinical exomes/genomes for a list of known variants in genes associated with medically important conditions	It is preferable to use a targeted approach to avoid unsolicited findings, and genomic screening is not specifically advocated
Patients cannot opt-out of genomic screening; it is the responsibility of the clinical team to provide appropriate pre- and post-test counselling	Guidelines for informed consent need to be developed, but patients' claims to a right not to know do not automatically over-ride professional responsibilities
The genomes of minors should be screened for variants offering clinical utility to their parents	Guidelines for testing minors need to be developed relating to what "unsolicited information" should be disclosed

Figures

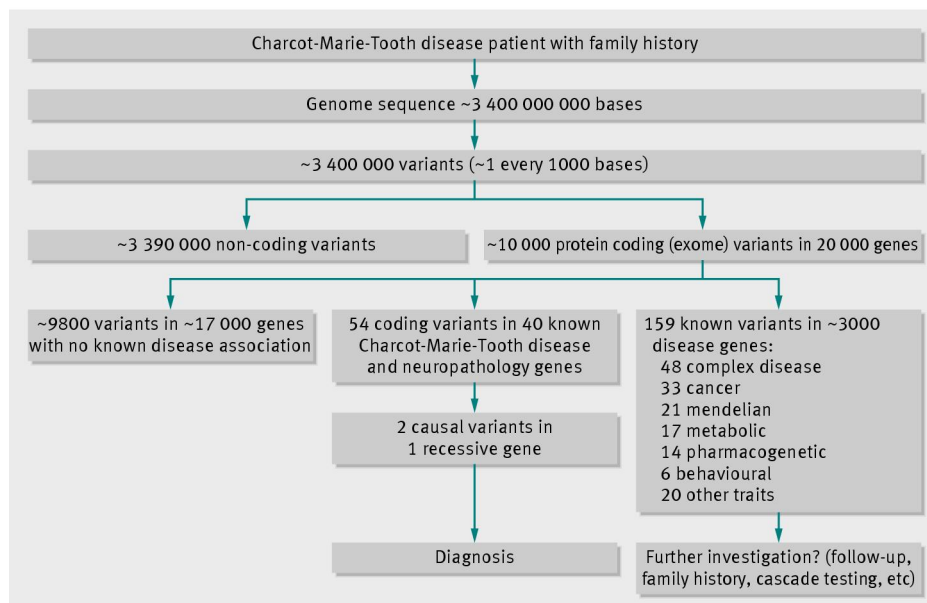


Fig 1 Example of patient genome analysis.¹⁰ A man was investigated to find the cause of his familial Charcot-Marie-Tooth syndrome, for which a genetic diagnosis was subsequently achieved. But what should be done with information on other variants relating to different diseases and traits?

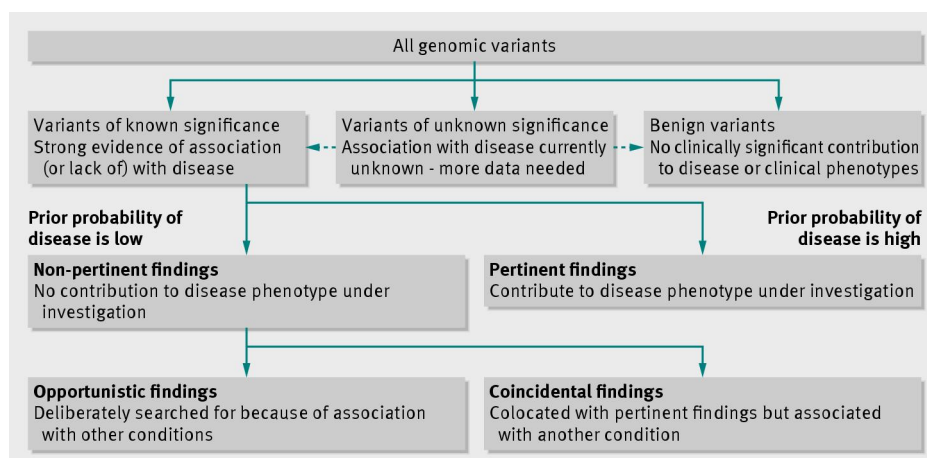


Fig 2 Categorisation of variants. By restricting clinical genomic analysis to a targeted approach based on the clinical question, it is possible to maximise pertinent findings and minimise non-pertinent ones. Only coincidental findings cannot be avoided using this strategy, such as a large deletion in a developmentally delayed child affecting genes for both a neurodevelopmental disorder and predisposition to cancer