Immunology & Cell Biology 2021; **99**: 168–176
www.wileyonlinelibrary.com/journal/icb

SPECIAL FEATURE REVIEW

The future is now? Clinical and translational aspects of "Omics" technologies

Gemma L D'Adamo[1], James T Widdop[1] & Edward M Giles[1,2]

1 Centre for Innate Immunity and Infectious Disease, Hudson Institute of Medical Research, Clayton, VIC, Australia
2 Department of Paediatrics, Monash University, Clayton, VIC, Australia

Keywords
Artificial intelligence, genomics, machine learning, microbiome, translational immunology

Corresponding Author
Edward M Giles, Department of Paediatrics, Monash University, 246 Clayton Road, Clayton, VIC 3168, Australia.
E-mail: Edward.giles@monash.edu

Received 29 July 2020; Revised 9 September 2020; Accepted 9 September 2020

doi: 10.1111/imcb.12404

Immunology & Cell Biology 2021; **99**: 168–176

Abstract

Big data has become a central part of medical research, as well as modern life generally. "Omics" technologies include genomics, proteomics, microbiomics and increasingly other omics. These have been driven by rapid advances in laboratory techniques and equipment. Crucially, improved information handling capabilities have allowed concepts such as artificial intelligence and machine learning to enter the research world. The COVID-19 pandemic has shown how quickly information can be generated and analyzed using such approaches, but also showed its limitations. This review will look at how "omics" has begun to be translated into clinical practice. While there appears almost limitless potential in using big data for "precision" or "personalized" medicine, the reality is that this remains largely aspirational. Oncology is the only field of medicine that is widely adopting such technologies, and even in this field uptake is irregular. There are practical and ethical reasons for this lack of translation of increasingly affordable techniques into the clinic. Undoubtedly, there will be increasing use of large data sets from traditional (e.g. tumor samples, patient genomics) and nontraditional (e.g. smartphone) sources. It is perhaps the greatest challenge of the health-care sector over the coming decade to integrate these resources in an effective, practical and ethical way.

INTRODUCTION

Over the last two decades, there has been a huge expansion in the use of "omics"-based approaches in medical research, which incorporate technologies characterizing various biological products, including genes (genomics), messenger RNA (transcriptomics), protein (proteomics) and metabolites (metabolomics) in biological samples.[1] Technological advancements in laboratory-based protocols, data storage and bioinformatic capabilities have ensured that it is now possible to generate huge amounts of "omics" data in a cost- and time-efficient manner. This has been demonstrated by the vast amount of COVID-19 research data produced in a matter of months. However, the general clinical uptake of these technologies so far remains restricted to a select few areas. Idealistic terms such as "precision medicine" and "pharmacogenomics" are commonly used, but remain inconsistently defined and largely aspirational.

The integration of big data and machine learning (ML) into "omics" unlocks the ability to collect, process and integrate huge amounts of health-related information in an extremely fast and efficient manner. ML models, which in essence apply the principles of artificial intelligence (AI), can "learn" from patterns in data sets, and can therefore accurately identify complex associations between variables which may not be possible by scientists and simple computer algorithms. The extremely varied nature of the data, in terms of type and source, has necessitated technological advancements to aid with fast, cost-effective data analysis in the research sphere. Despite this, there remain significant barriers to the lack of utilization of "omics" in clinical practice. These include lack of clinician knowledge (including in definitions—see Table 1) and ethical considerations.[2] In particular, there

are potential issues in overdiagnosis and identification of abnormalities which pose no clinically significant risk.[1]

The focus of this review will be to explore the areas of clinical medicine where omics and big data are already shaping clinical management, or are on the cusp of doing so. First, we will focus on the oncology, which has seen the largest uptake of these techniques. Then, we will explore the integration of multiomics in complex diseases. Owing to the significant burden they impose on Western health-care systems, the implementation of advanced diagnostics, prognostication and data-guided treatment regimens should allow improvements in individual patient care and may decrease health-care burden. In addition, interest in the microbiome as a "new" large data set in human health will be explored. Finally, we will address some of the challenges ahead for clinicians and researchers, both practical and ethical, in this fast-changing landscape.

TUMOR (GEN)OMICS

The main implementation of "omics" in clinical practice thus far has been in oncology genomics. Over the last decade, well-defined genomic markers have enabled robust screening and guided therapeutic management.[3] The analysis of somatic mutations in tumors has demonstrated the power of this technology to predict treatment response while minimizing risks of adverse effects. Precision oncology, which centers around the selection of immunotherapies and chemotherapeutic agents, based on tumor molecular profiling, provides an avenue for individualized care (Figure 1). For example, analyzing epidermal growth factor receptor mutations in non-small-cell lung cancers can guide immunotherapy.[4] However, there are multiple challenges on both the individual and systems levels which are inhibiting the integration of these approaches into everyday care. Data sets remain incomplete; in the non-small-cell lung cancers example patient cohorts often only express the most common mutations.[5,6] In addition, the interindividual variability in treatment response is still significant. This difference may be a result of patient-intrinsic factors, including the physiological distribution of disease, underlying genetics and epigenetic background of the patient,[4,7] highlighting the potential need for integration of other omics data into decision making.

TUMOR PHARMACO-GENOMICS

Undoubtedly, the future of precision medicine will center around linking multiomic patient profiles, in order to guide disease management and treatment regimens. For example, there are resources that analyze large data sets of both patients and drugs, such as the CDRscan,[8] which are already showing the great promise of these methods in oncology. CDRscan utilizes a novel deep learning model that predicts anticancer drug responsiveness based on large-scale drug screening data, by combining the genomic profiles of 787 human cancer cell lines with the structural profiles of 244 drugs. Such tools act to integrate information pertaining to patient tumor molecular profiles with the molecular fingerprints of the drugs, by merging them *in silico* to create a "virtual docking" model. The power of such resources is that they enable the selection of antitumor (and potentially other) agents based off individual genomic profiles, and provide an avenue for feasibly predicting effective and safe drug repurposing opportunities. This should streamline patient-tailored drug selection as a support system for clinical decision making.

The integration of multiomics data with physiological profiling from well-characterized cohorts provides an accessible avenue for exploring new drug targets. This work may also enable characterization of parameters that define drug responders from nonresponders and identify patients who are unable to process prodrugs.[9] The potential benefits of pharmacogenomic screening are highlighted by genetic testing for polymorphisms associated with the cytochrome P450 family of enzymes. These enzymes are extremely important for catalyzing a

Table 1. Glossary of important terms.

Term	Definition
Precision medicine	Medical care tailored to individuals based on particular characteristics or phenotypes. Often used interchangeably with the older term "personalized" medicine.
"Omics"	The use of biotechnology to interrogate biological function at a particular "level." Can describe any combination of genomics, transcriptomics, proteomics, metabolomics.
Big data	Large and diverse data sets that lend themselves to be analyzed computationally.
Pharmacogenomics	The study of determining an individual's drug response based on his/her genetic profile.
Machine learning	Complex computer algorithms that are able to self-improve following the input of relevant data.
Artificial intelligence	Computer systems capable of performing tasks generally requiring humans, such as decision making.

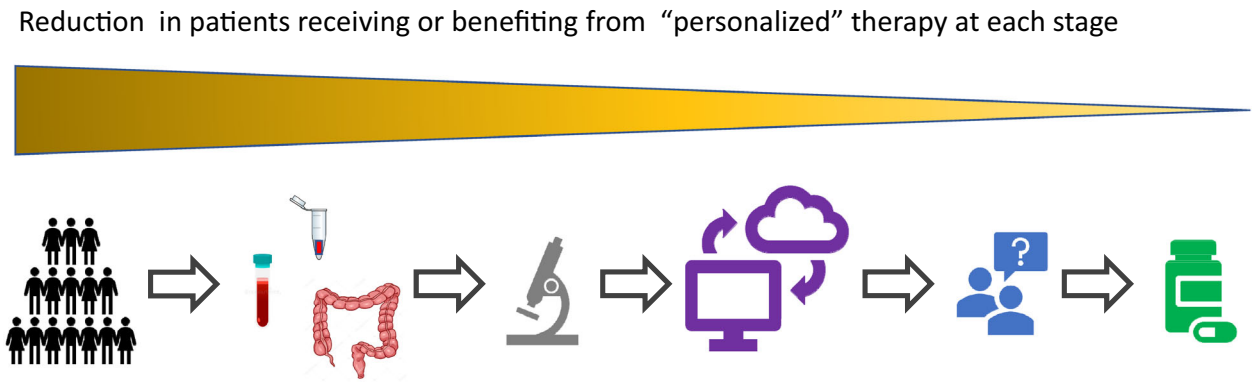

Figure 1. Current pipeline of "omics" to develop precision medicines. The schematic is described as possible pipeline for development of precision medicine based on large data sets. Data (genomic and other data from patients and/or their area of disease, e.g. tumor) are analyzed. This analysis is compared with large publicly available data sets. This output is discussed by the clinician in conjunction with other clinical data available. Decisions can then be made on appropriate treatments or trials for each individual. This current model "loses" patients at each step, because of lack of sample availability, access to particular analysis, poor publicly available data sets, other confounding medical conditions or availability of trials or safe appropriate targeted therapies.

significant proportion (70–80%) of medications currently used for clinical purposes. They also demonstrate high levels of variability at the genetic level, making screening important in the prediction of both potential adverse reactions and clinical response to drugs.[10] For example, screening patients for CYP variants to predict likelihood of toxicity has become standard of care in prescribing fluropyrimidines in the treatment of multiple cancers.[11]

In addition, pharmacogenomics potentially has an important role to play in guiding treatment selection by predicting potential toxicity to therapeutic agents[11,12] and directing medication dosing.[13] Variants in drug targets affecting pharmacodynamic interactions have been utilized clinically in predicting response to particular therapies beyond oncology. For example, a drug used in cystic fibrosis, ivacaftor, is effective in patients that have a *specific* mutation (phe508del) in the CFTR target chloride channel.[14]

Despite the great potential pharmacogenomics presents, there has been a lack of implementation in clinical practice so far, largely because of the scarcity of clear, decisive and well-reviewed guidelines. Collaborative efforts between clinicians, scientists and bioinformaticians will be essential in generating these resources.[15] Nevertheless, the scope for pharmacogenomics in improving health-related processes is immense in terms of reducing diagnostic and treatment-related errors, eliminating redundant tests and improving the distribution of scarce health resources.

PATIENT (GEN)OMICS IN ONCOLOGY AND ELSEWHERE

Following on from the expanding field of tumor-based sequencing informing clinical management decisions, there have also been efforts made in the clinic to sequence patient DNA, rather than the tumor genome.[16,17] Unlike sequencing the tumor genome for driver mutations to predict response to therapies, sequencing the patient DNA aims to identify germline (inherited) mutations that would be present in all cells. One of the biggest areas of uptake of genomic testing in the clinic has been in the context of screening patients, to either predict risk of developing disease or patient response to medications. These genomic methods are currently being used in oncology,[18] but also in others fields such as cardiology[19] and neurology.[20] The method of identifying mutations in these regions is also changing from techniques that recognized known DNA mutations to next-generation sequencing approaches.[21]

Initially, patient DNA was collected using arrays of known single-nucleotide polymorphisms within genes that conferred increased risk. Now, common genomic tests sequence a gene or multiple genes of interest and are able to detect a wider range of genomic variation (single-nucleotide polymorphisms, insertions, deletions, translocations) within these regions. Currently, the most common method of targeted gene sequencing is using multigene panels for specific diseases. For example, a patient who has a strong family history of breast cancer may receive a blood test that sequences a panel of genes known to be associated with breast cancer.[22] However, with next-generation sequencing becoming increasingly efficient and more widely available, patients are now able to have their entire genome sequenced. However, these high-throughput methods create a large increase in the amount of data returned to the clinician for interpretation. Ethical considerations pertaining to how

this information is communicated to the patient are yet to be resolved.[23] A significant concern is that full-genome sequencing is already leading to incidental findings of pathogenic variants in patients' genome, and without other clinical data such as family history, or other omics data such as expression levels (transcriptomics, proteomics), results are likely to be flawed in directing clinical decision making.

In addition, most recommendations in clinical practice are restricted to genes which represent a small proportion of the genome that codes for protein. Clearly, there is untapped potential in determining the significance of variants in noncoding sequences (such as promoter or enhancer regions of DNA) but also ethical considerations of sequencing this material without clinically actionable interventions available.

One of the more well-known and controversial examples of clinical genomic sequencing is that of the BRCA1/2 gene mutations. The BRCA1 and BRCA2 genes both encode for tumor suppressor proteins, whose variants have been linked with breast and ovarian cancer, together named hereditary breast and ovarian cancer syndrome. Patients with pathogenic BRCA mutations are likely to develop cancer earlier and have more aggressive phenotypes than other forms of breast cancer.[24] A recent study reported a lifetime risk of developing breast cancer of 72% among patients with BRCA1 and 69% of patients with harmful BRCA2 mutations develop breast cancer.[18] In the general population pathogenic variants of the BRCA gene is low.[25] In the United States, BRCA testing is only recommended following a medical history-based questionnaire and consultation with a genetic counselor.[26] Thus, with the increasing availability of DNA sequencing with next-generation sequencing technology, there have been suggestions to implement this into a population-based screening approach, with one study showing this to be a clinically useful cost-effective option.[27]

Guidelines have been developed to address some of these clinical genomic issues. Currently the American College of Molecular Genetics has identified pathogenic variants in 59 different genes,[28] and recommends reporting back to the patient regardless of the indication for testing.

THE FUTURE IS . . . NEARLY NOW—ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING

As these large clinical data sets expand rapidly, it becomes critical to find new and better ways to manage this information. The power of AI and ML for doing this has been recognized in many fields. Over the last decade, the development of cloud computing systems and ML pipelines has opened the door to significantly improve our ability to perform noninvasive disease diagnostics, prognostication and monitoring using big data sets derived from multiple sources. ML is needed to cope with the vast volume and significant heterogeneity that characterizes many health-care-related data sets. However, before sophisticated analyses can be performed for clinical purposes, the algorithms must be "trained" using well-organized, structured and linked data sets. This training step is essential for developing accurate and efficient pipelines. Databases, such as the Cancer Dependency Map,[29] provide a resource for these purposes, as it contains linked molecular profiles, drug responses and genetic viability data on more than 1000 cancer cell lines. In addition, the AACR Project GENIE[30] and ASCO CancerLinQ[31] are also collecting genomic profiles linked with clinical data for thousands of patients, which will provide a valuable resource for the field. The importance of large, complete data sets such as these cannot be understated in the development and training of accurate ML algorithms, which are fundamentally important for their translatability to the health-care sector.

Undoubtedly, the future of big data in the health-care sector will also incorporate home health monitoring, via smartphones and smart health-care trackers, which track multiple physiological and biometric parameters, to enable advancements in individualized care and personalized management for patients (Figure 2). Continual noninvasive monitoring is important for

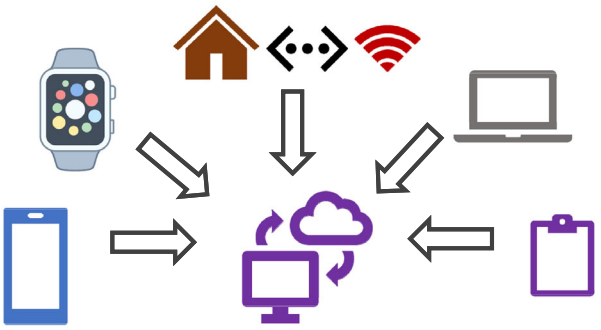

Figure 2. Use of unconventional data sets for health care. Devices such as smartphones and smartwatches are increasing capable of collecting health-related data (e.g. heart rate, movement information). Other large data sets from nontraditional setting may include internet-enabled home devices (e.g. refrigerator, lighting), as well as browser histories or even traditional medical records being centrally stored. The provision of such data for central analysis offers enormous opportunities for understanding both individual and population health. Such data have equally enormous practical and ethical challenges in their use.

establishing an individual's baseline and then deviations from that baseline can be detected, even using smart electronic devices, which could either provide health advice at home or encourage the person to seek medical attention. A number of prototypes have been designed and investigated for a far-reaching variety of medical conditions, including diabetes management,[32] detection of atrial fibrillation,[33] blood cholesterol monitoring,[34] early detection of Parkinson's disease,[35] self-adherence and early warning of heart attack.[36] Speech-driven home assistants have also been used to detect agonal breathing, an audible biomarker that is an early sign of cardiac arrest.[37] Establishing the infrastructure to enable these analyses to be performed in a streamlined and efficient manner will be essential. These resources provide the foundation for noninvasive at-home medical management, which not only allows for more detailed and frequent patient monitoring but also reduces the potential load placed on physicians to do this work.

AI IN DIAGNOSTICS

The ability of AI systems to improve upon detection rates for diseases has been investigated.[38,39] McKinney *et al.* compared the ability of a deep-learning AI model with independent radiologists in the accurate detection of early stage breast malignancy. They utilized two large intercontinental cohorts, derived from the United States and the UK, and performed retrospective clinical comparisons to evaluate the predictive power of the AI system used. Interestingly, they demonstrated the AI system to have a superior ability to predict early stage malignancy in both cohorts, which differed in both population characteristics and screening practices. Therefore, they suggested that utilization of AI systems may enable detection of malignancy before standard care would. Complex detection models such as this could be used in conjunction with experienced clinicians, to flag suspicious regions of tissue for expert review, triage patients for urgent review if a "high-risk" lesion is identified and dismiss more "low-risk" cases. In addition, systems such as these may provide valuable resources for delivering sustainable screening programs in countries with workforce shortages in the health-care sector. Therefore, the development of ML systems to advance diagnostics may not only improve the detection accuracy, but also reduce the load on the already strained public health system.

AI FOR PROGNOSTICS

Deep learning AI and ML models not only have a role to play in diagnostics, but also allow for the integration of complex multiomics data sets for disease prognostication. The potential benefit of these integrative analyses has become evident for some life-limiting conditions, such as hepatocellular carcinoma, which has a 5-year survival rate of about 30%.[40] The complexity of hepatocellular carcinoma stems from the high level of intradisease heterogeneity, with multiple complex etiological factors making prognostic predictions extremely challenging. However, studies have shown that the integration of multiple omics data sets can stratify patients into disease subgroups based off molecular profiling, which aid treatment guidance and prognostication. This was demonstrated by assessing one database, which was able to be successfully grouped via unsupervised analyses. These results were then validated in five independent cohorts, containing different omic's data sets (microRNA or messenger RNA and DNA methylation). Overall, this study demonstrated that the high-powered computational rigor able to be achieved using ML algorithms has the ability to identify discordant molecular disease subtypes in a robust manner. Similar work has been performed in other disease processes displaying high levels of heterogeneity, with promising results.[41–43] The complexity of these conditions means that there are limitations associated with the use of single omics data in disease classification, and therefore, more clinically significant disease insights will stem from integration of multiomics data sets.

AI TO THE PEOPLE? IMPROVING ACCESS TO HEALTH CARE BY TECHNOLOGY

Another advantage in the use of such technology is access to expertise without the need for direct clinician contact. For example, the diagnosis of skin malignancies traditionally follows the path of clinical examination, dermatoscopic analysis and then histological examination. The distinction between malignant and benign skin lesions can be challenging to even experienced dermatologists, because of subtle differences in the appearance of lesions. Deep convolutional neural networks have previously been shown to have good power to perform classification tasks on objects with fine-grained differences in appearance. Esteva *et al.*[44] aimed to investigate the power of these approaches for classification of skin lesions. This AI model was shown to perform on par with clinician assessment, and represents the possibility of extending diagnostics beyond the clinic. Another example is in diabetes, where retinopathy[45] was able to be assessed for by computer-aided detection software with high levels of accuracy, comparable to clinicians. This has the potential to expand into a plethora of other medical specialties involved in complex

disease management, such as ophthalmology, otolaryngology and radiology, thereby reducing the load on the health care system and improving the provision of medical care to the wider community, particularly in regional and remote areas.

MICROBIOMICS—A NEWISH FRONTIER

Another omics type that has exponentially increased in the last two decades is study of the human microbiome. Compared with our 20 000 genes, the human gut is estimated to harbor many millions of microbial genes and offers a huge amount of genetic diversity which is in theory dynamic and changeable, making it a very attractive target for therapeutic manipulation in an increasing number of diseases as diverse as cancer, autism and inflammatory bowel disease (IBD).[46–48]

There has been an explosion in research in the microbiome in cancer research. Perhaps the most significant practical discoveries have been on response to therapy. The microbiome, and its interindividual diversity, is believed to be an important factor in the development of cancer, and the diverse response to treatments (discussed above). For example, anti-PD-1 therapy in melanoma and epithelial cell cancer patients is strongly influenced by the gut microbiome.[49,50] Indeed, metagenomic analysis of patient stool samples of anti-PD-1 responders has identified a bacterial species that could improve the efficacy of the anti-PD-1 therapy in previous nonresponders. It is only through the combination of bioinformatic analysis of both microbiome data sets with other clinical data sets that other potential modifiers of treatment response will be found.

IBD, a chronic inflammatory disease affecting the gastrointestinal tract, has been an obvious choice for studying the microbiome. Omics technologies have given us insight into the complex etiology of the condition, which likely involves genetic susceptibility and changes to the microbiota.[46] Recent efforts have employed longitudinal approaches and combined omics technologies on the host and microbial side which have given us detailed insight into changes in microbial host relationships during health and IBD.[51,52] However, even with these efforts a lack of sequencing depth precludes our ability to determine exactly which strains of microbes are present in each patient. It is well understood with pathogens, but only recently recognized in microbiome research, that closely related species may be functionally extremely different.[53] This has large implications for limiting translation of therapeutic approaches in IBD and other microbiome research into the clinic.

Indeed, IBD is a good example of a complex disease, with a polygenic background and strong but poorly understood environmental influences that require understanding enormous complex and diverse data sets to understand the disease. There have been preliminary attempts at developing such a "multiomics" approach to IBD patients.[51,54,55] In this proposed model, described in Figure 3, data from multiple sources from each individual would be collected, processed and analyzed. Novel and evolving informatic techniques would be used to determine individualized targets for therapy. There are clear advantages to this model, where patients would receive treatments with a much higher chance of success, and reduction in side effects and wasted expense of ineffective therapies. While this approach has yet to show clinical benefit, it is an active approach being explored in many chronic complex diseases including Type 2 diabetes, asthma and schizophrenia.

CHALLENGES FOR IMPLEMENTATION AND ETHICAL CONSIDERATIONS

Despite the great promise presented by the integration of multiomics, big data, AI and ML into the sphere of clinical medicine, there are limitations and ethical considerations currently inhibiting its widespread implementation. Primarily, there is a fundamental lack of consensus over the clinically relevant genetic variants to focus on[18] and a lack of recognition by clinicians, hospitals and funding bodies of the potential for genomic-guided care to improve patient outcomes. These factors, coupled with concerns over the cost of omics in medicine and the general resistance to change in health-care systems, mean that significant bodies of evidence are required before these techniques will be adopted more widely.

Alongside the logistic concerns surrounding implementation of these technologies are the equally important ethical considerations.[56] Issues surrounding informed consent for testing and how the sizeable amount of data generated for each individual will be managed, distributed and reported upon remain of paramount importance. Consent is extremely challenging, because of the significant number of genetic loci being assessed and their potential influence on disease development.

Because of this complexity and diversity, prescreen counseling of patients on the scope of potential findings is likely not possible. The volume of data generated by whole-genome sequencing exceeds levels that could be reasonably assessed by clinicians and presented to patients, which conflicts with general expectations placed upon medical professionals. However, in order to ensure

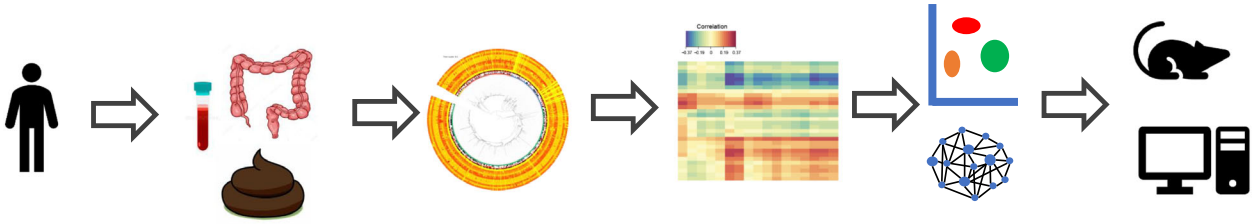

Figure 3. Proposed pipeline of omics for precision medicine in chronic disease. A patient with inflammatory bowel disease is an example, where samples (e.g. blood, stool, biopsy) are simultaneously analyzed, forming large data sets that can be simultaneously analyzed. Patterns are determined using validated informatic techniques. Individualized targets for therapy can be tested by animal models, *in silica* analysis, organoid and/or "organ-on-a-chip." For example, patient-specific intestinal organoids have been used to confirm response to therapy in cystic fibrosis patients.

that patients retain the most important information, it is important that they are not encumbered with large volumes of potentially insignificant testing results. This issue is clouded by limited familiarity among many clinicians with genetic testing results and differing perspectives among clinicians on the importance of returning screening results. While some clinicians see value in relaying results to patients to enable screening of family members, others restrict reporting results to those which influence immediate patient management. There are also implications for long-term follow-up of patients when screening results are not immediately actionable, and a need to consider whether the required services are available, including counseling providers and the financial resources required to meet patient needs. This is particularly important in the current climate with costs associated with whole-genome sequencing declining, while those associated with patient follow-up, interventions and medical evaluation are not. The complexity of this situation is also such that patients may become burdened, with far-reaching impacts on familial relationships and the ability of patients to secure health insurance moving forward. Therefore, there is a need for standardized assessment and reporting criteria to be established for genetic screening before its widespread implementation is possible.

These ethical and practical challenges must be balanced against the enormous potential across all the disciplines of clinical medicine, from diagnostics, management and prognostics. It will only be through clear guidance from policy makers, based on better evidence, and clinician education that will inform the pathway of increased clinical omics usage over the next decade. The recent broadening of criteria for government-funded whole-genome sequencing in children in Australia, coupled with a rollout of genetic education through various methods, is an excellent example of such a change.

SUMMARY

During the current COVID-19 crisis, technology has provided an opportunity for providing ongoing health care in novel ways that would have been inconceivable a decade ago. This has shown that in a crisis, the health-care system can adapt to new technologies at speed. In the research field, we are overwhelmed with data on SARS-CoV-2 journal articles, with novel omics techniques and informatics providing much of the ability for such rapid data acquisition and analysis. However, the rush to develop a rapid vaccine remains challenged by practical, ethic as well as technical and immunological difficulties. This mirrors the challenges faced by clinicians in introducing new omics technologies across almost all areas of health care.

Indeed, the widespread use of big data in business, media, politics and research has struggled to penetrate clinical medicine. The potential clinical benefit of this technology is enormous. While cost is a factor, there is also potential for AI to provide health care in a more affordable way, especially in areas with limited skilled clinicians. Practical considerations, including ethical and informatic issues, remain a significant barrier to widespread uptake of omics in medicine. It is the challenge of policymakers, funders and clinicians to navigate this new frontier in health care.

ACKNOWLEDGMENTS

The authors thank Dr Jo Lundy for her insights into the use of omics in clinical oncology.

AUTHOR CONTRIBUTION

Edward M Giles: Writing-original draft; Writing-review & editing. **Gemma L D'Adamo:** Writing-original draft; Writing-

review & editing. **James T Widdop:** Writing-original draft; Writing-review & editing.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

REFERENCES

1. Shilo S, Rossman H, Segal E. Axes of a revolution: challenges and promises of big data in healthcare. *Nat Med* 2020; **26**: 29–38.
2. Goecks J, Jalili V, Heiser LM, Gray JW. How Machine Learning Will Transform Biomedicine. *Cell* 2020; **181**: 92–101.
3. Malone ER, Oliva M, Sabatini PJB, Stockley TL, Siu LL. Molecular profiling for precision cancer therapies. *Genome Med* 2020; **12**: 8.
4. Kobayashi Y, Mitsudomi T. Not all epidermal growth factor receptor mutations in lung cancer are created equal: Perspectives for individualized treatment strategy. *Cancer Sci* 2016; **107**: 1179–1186.
5. Takeuchi K, Soda M, Togashi Y, *et al.* RET, ROS1 and ALK fusions in lung cancer. *Nat Med* 2012; **18**: 378–381.
6. Zhuo M, Zheng Q, Zhao J, *et al.* Survival difference between EGFR Del19 and L858R mutant advanced non-small cell lung cancer patients receiving gefitinib: a propensity score matching analysis. *Chin J Cancer Res* 2017; **29**: 553–560.
7. Rotow J, Bivona TG. Understanding and targeting resistance mechanisms in NSCLC. *Nat Rev Cancer* 2017; **17**: 637–658.
8. Chang Y, Park H, Yang HJ, *et al.* Cancer Drug response profile scan (CDRscan): a deep learning model that predicts drug effectiveness from cancer genomic signature. *Sci Rep* 2018; **8**: 8857.
9. Gill D, Georgakis MK, Koskeridis F, *et al.* Use of Genetic Variants Related to Antihypertensive Drugs to Inform on Efficacy and Side Effects. *Circulation* 2019; **140**: 270–279.
10. Westergaard N, Sogaard Nielsen R, Jorgensen S, Vermehren C. Drug use in Denmark for drugs having pharmacogenomics (PGx) based dosing guidelines from CPIC or DPWG for CYP2D6 and CYP2C19 drug-gene Pairs: Perspectives for introducing PGx test to polypharmacy patients. *J Pers Med* 2020; **10**: 3.
11. Henricks LM, Lunenburg C, de Man FM, *et al.* DPYD genotype-guided dose individualisation of fluoropyrimidine therapy in patients with cancer: a prospective safety analysis. *Lancet Oncol* 2018; **19**: 1459–1467.
12. Deenen MJ, Meulendijks D, Cats A, *et al.* Upfront Genotyping of DPYD*2A to Individualize Fluoropyrimidine Therapy: A Safety and Cost Analysis. *J Clin Oncol* 2016; **34**: 227–234.
13. Jorgensen AL, Prince C, Fitzgerald G, *et al.* Implementation of genotype-guided dosing of warfarin with point-of-care genetic testing in three UK clinics: a matched cohort study. *BMC Med* 2019; **17**: 76.
14. Wainwright CE, Elborn JS, Ramsey BW, *et al.* Lumacaftor-Ivacaftor in Patients with Cystic Fibrosis Homozygous for Phe508del CFTR. *N Engl J Med* 2015; **373**: 220–231.
15. Ji Y, Si Y, McMillin GA, Lyon E. Clinical pharmacogenomics testing in the era of next generation sequencing: challenges and opportunities for precision medicine. *Expert Rev Mol Diagn* 2018; **18**: 411–421.
16. Hart MR, Biesecker BB, Blout CL, *et al.* Secondary findings from clinical genomic sequencing: prevalence, patient perspectives, family history assessment, and health-care costs from a multisite study. *Genet Med* 2019; **21**: 1100–1110.
17. Manolio TA, Bult CJ, Chisholm RL, *et al.* Genomic medicine year in review: 2019. *Am J Hum Genet* 2019; **105**: 1072–1075.
18. Kuchenbaecker KB, Hopper JL, Barnes DR, *et al.* Risks of breast, ovarian, and contralateral breast cancer for BRCA1 and BRCA2 mutation carriers. *JAMA* 2017; **317**: 2402–2416.
19. Palomaki GE, Melillo S, Neveux L, *et al.* Use of genomic profiling to assess risk for cardiovascular disease and identify individualized prevention strategies–a targeted evidence-based review. *Genet Med* 2010; **12**: 772–784.
20. Poduri A, Sheidley BR, Shostak S, Ottman R. Genetic testing in the epilepsies-developments and dilemmas. *Nat Rev Neurol* 2014; **10**: 293–299.
21. Tung N, Battelli C, Allen B, *et al.* Frequency of mutations in individuals with breast cancer referred for BRCA1 and BRCA2 testing using next-generation sequencing with a 25-gene panel. *Cancer* 2015; **121**: 25–33.
22. Easton DF, Pharoah PD, Antoniou AC, *et al.* Gene-panel sequencing and the prediction of breast-cancer risk. *N Engl J Med* 2015; **372**: 2243–2257.
23. Solomon BD. Incidentalomas in genomics and radiology. *N Engl J Med* 2014; **370**: 988–990.
24. Baretta Z, Mocellin S, Goldin E, Olopade OI, Huo D. Effect of BRCA germline mutations on breast cancer prognosis: A systematic review and meta-analysis. *Medicine (Baltimore)* 2016; **95**: e4975.
25. Eng C, Hampel H, de la Chapelle A. Genetic testing for cancer predisposition. *Annu Rev Med* 2001; **52**: 371–400.
26. Moyer VA, Force USPST. Risk assessment, genetic counseling, and genetic testing for BRCA-related cancer in women: U.S. Preventive Services Task Force recommendation statement. *Ann Intern Med* 2014; **160**: 271–281.
27. Petelin L, Hossack L, Shanahan M, *et al.* Cost-effectiveness of long-term clinical management of BRCA pathogenic variant carriers. *Genet Med* 2020; **22**: 831–839.
28. Manolio TA, Rowley R, Williams MS, *et al.* Opportunities, resources, and techniques for implementing genomics in clinical care. *Lancet* 2019; **394**: 511–520.
29. Tsherniak A, Vazquez F, Montgomery PG, *et al.* Defining a Cancer Dependency Map. *Cell* 2017; **170**: 564–576 e516.
30. Consortium APG. AACR project GENIE: Powering precision medicine through an international consortium. *Cancer Discov* 2017; **7**: 818–831.

31. Miller RS, Wong JL. Using oncology real-world evidence for quality improvement and discovery: the case for ASCO's CancerLinQ. *Future Oncol* 2018; **14**: 5–8.
32. Zheng T, Xie W, Xu L, *et al*. A machine learning-based framework to identify type 2 diabetes through electronic health records. *Int J Med Inform* 2017; **97**: 120–127.
33. Bumgarner JM, Lambert CT, Hussein AA, *et al*. Smartwatch Algorithm for Automated Detection of Atrial Fibrillation. *J Am Coll Cardiol* 2018; **71**: 2381–2388.
34. Fu Y, Guo J. Blood Cholesterol Monitoring With Smartphone as Miniaturized Electrochemical Analyzer for Cardiovascular Disease Prevention. *IEEE Trans Biomed Circuits Syst* 2018; **12**: 784–790.
35. Lonini L, Dai A, Shawen N, *et al*. Wearable sensors for Parkinson's disease: which data are worth collecting for training symptom detection models. *NPJ Digit Med* 2018; **1**: 64.
36. Sahoo PK, Thakkar HK, Lee MY. A Cardiac Early Warning System with Multi Channel SCG and ECG Monitoring for Mobile Health. *Sensors (Basel)* 2017; **17**:711.
37. Chan J, Rea T, Gollakota S, Sunshine JE. Contactless cardiac arrest detection using smart devices. *NPJ Digit Med* 2019; **2**: 52.
38. Wu N, Phang J, Park J, *et al*. Deep neural networks improve radiologists' performance in breast cancer screening. *IEEE Trans Med Imaging* 2020; **39**: 1184–1194.
39. McKinney SM, Sieniek M, Godbole V, *et al*. International evaluation of an AI system for breast cancer screening. *Nature* 2020; **577**: 89–94.
40. Chaudhary K, Poirion OB, Lu L, Garmire LX. Deep learning-based multi-omics integration robustly predicts survival in liver cancer. *Clin Cancer Res* 2018; **24**: 1248–1259.
41. Asada K, Joutard S, *et al*. Uncovering prognosis-related genes and pathways by multi-omics analysis in lung cancer. *Biomolecules* 2020; **10**:524.
42. Chang P, Grinband J, Weinberg BD, *et al*. Deep-Learning Convolutional Neural Networks Accurately Classify Genetic Mutations in Gliomas. *AJNR Am J Neuroradiol* 2018; **39**: 1201–1207.
43. Mobadersany P, Yousefi S, Amgad M, *et al*. Predicting cancer outcomes from histology and genomics using convolutional networks. *Proc Natl Acad Sci USA* 2018; **115**: E2970–E2979.
44. Esteva A, Kuprel B, Novoa RA, *et al*. Dermatologist-level classification of skin cancer with deep neural networks. *Nature* 2017; **542**: 115–118.
45. Gulshan V, Peng L, Coram M, *et al*. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. *JAMA* 2016; **316**: 2402–2410.
46. McIlroy J, Ianiro G, Mukhopadhya I, Hansen R, Hold GL. Review article: the gut microbiome in inflammatory bowel disease-avenues for microbial management. *Aliment Pharmacol Ther* 2018; **47**: 26–42.
47. Helmink BA, Khan MAW, Hermann A, Gopalakrishnan V, Wargo JA. The microbiome, cancer, and cancer therapy. *Nat Med* 2019; **25**: 377–388.
48. Yang Y, Tian J, Yang B. Targeting gut microbiome: A novel and potential therapy for autism. *Life Sci* 2018; **194**: 111–119.
49. Routy B, Le Chatelier E, Derosa L, *et al*. Gut microbiome influences efficacy of PD-1-based immunotherapy against epithelial tumors. *Science* 2018; **359**: 91–97.
50. Gopalakrishnan V, Spencer CN, Nezi L, *et al*. Gut microbiome modulates response to anti-PD-1 immunotherapy in melanoma patients. *Science* 2018; **359**: 97–103.
51. Lloyd-Price J, Arze C, Ananthakrishnan AN, *et al*. Multi-omics of the gut microbial ecosystem in inflammatory bowel diseases. *Nature* 2019; **569**: 655–662.
52. Poyet M, Groussin M, Gibbons SM, *et al*. A library of human gut bacterial isolates paired with longitudinal multiomics data enables mechanistic microbiome research. *Nat Med* 2019; **25**: 1442–1452.
53. Kamada N, Chen GY, Inohara N, Nunez G. Control of pathogens and pathobionts by the gut microbiota. *Nat Immunol* 2013; **14**: 685–690.
54. Ananthakrishnan AN, Luo C, Yajnik V, *et al*. Gut microbiome function predicts response to anti-integrin biologic therapy in inflammatory bowel diseases. *Cell Host Microbe* 2017; **21**: 603–610 e603.
55. Kugathasan S, Denson LA, Walters TD, *et al*. Prediction of complicated disease course for children newly diagnosed with Crohn's disease: a multicentre inception cohort study. *Lancet* 2017; **389**: 1710–1718.
56. Sharp RR. Downsizing genomic medicine: approaching the ethical complexity of whole-genome sequencing by starting small. *Genet Med* 2011; **13**: 191–194.