Translational informatics: an industry perspective

Michael N Cantor

Correspondence to

Michael N Cantor, Pfizer, Inc, and Department of Medicine, 235 East 42nd Street, New York, NY 10017, USA; michael.cantor@pfizer.com

Received 6 September 2011
Accepted 12 December 2011
Published Online First
11 January 2012

ABSTRACT

Translational informatics (TI) is extremely important for the pharmaceutical industry, especially as the bar for regulatory approval of new medications is set higher and higher. This paper will explore three specific areas in the drug development lifecycle, from tools developed by precompetitive consortia to standardized clinical data collection to the effective delivery of medications using clinical decision support, in which TI has a major role to play. Advancing TI will require investment in new tools and algorithms, as well as ensuring that translational issues are addressed early in the design process of informatics projects, and also given higher weight in funding or publication decisions. Ultimately, the source of translational tools and differences between academia and industry are secondary, as long as they move towards the shared goal of improving health.

The translational aspects of biomedical research gained a great deal of prominence through the establishment of the National Institutes of Health's clinical translational science awards.¹ In industry, and particularly the pharmaceutical industry, however, the translational aspects of research have always been at the forefront, as moving from bench to bedside (or target to approved drug) is the major reason most pharmaceutical companies exist. Improving the bench to bedside connection has become even more important for the industry lately, as it faces pressures from patent expirations, the fundamentally higher bar for future drug approvals as many medical needs become at least partly met, and a decreasing tolerance of adverse events.² On the discovery side, there is also increasing pressure to take advantage of the dramatic developments in data generation in both the clinical and preclinical arenas, leading to demands for more sophisticated informatics solutions.

TI: SCOPE

For the purposes of this paper, we define translational informatics (TI) as data-driven approaches that advance the movement of a compound, device, or intervention from the laboratory into the clinic. TI is also not confined to bioinformatics or -omic sciences. Because TI has been extensively reviewed elsewhere,^{3–6} we will look at three pivotal aspects of TI: early stage research, focusing on the role of precompetitive consortia; across research and development, providing incentives for the adoption of standards; and post-approval, focusing on the role of healthcare IT. Basic science research occurs in both the academic and industry settings, and in many cases, particularly in the area of data and

information management, the needs of industry and academia are well aligned. Translational issues enter this divide, however, when one considers the lack of uptake of many informatics tools developed in either setting, often because of scalability, applicability, or operational issues. No matter where its origin, a tool that results in a publication and then gathers dust in an archive does little to advance the field or meet researchers' needs.

TI IN RESEARCH: PRECISION MEDICINE

Moving more discoveries from bench to bedside is essential if the pharmaceutical industry is to continue to exist. Current pipelines, while containing important new products with great potential to improve patients' lives, are still generally not robust enough to match the results of the past decades in terms of productivity and replacing products that lose patent protection.² A key to improving productivity is improved understanding and use of data, which is why TI plays such a major role in the industry's future. Naturally this goal aligns well with the drug discovery process, in which the area of focus is hypothesis generation. Translational approaches, however, could actually permit more hypothesis testing in earlier stages, particularly by feeding back clinically relevant data into the research process, allowing for the earlier dismissal of mechanisms and targets with low probabilities of advancement. While narrowing the field of compounds that advance, this approach would also lead to a higher likelihood of success at later stages, leading to a much more efficient development process.

Developing medicines that facilitate the move from empirical to precision medicine is a major strategic goal for many pharmaceutical companies. A concurrent informatics strategy is essential to achieve the goal of delivering safer and more effective drugs to the right patient populations. Not only is the focus on -omics data to derive or confirm treatment or safety-related biomarkers, but also clinical data for identifying the cohorts of patients who can then be stratified along their likelihood of response (or adverse event) to a particular medicine, as well as the longitudinal data to evaluate the effectiveness and value of targeted medications over more empiric therapies.^{7–9}

PRECOMPETITIVE CONSORTIA

As the data needs of the industry grow and resources become more constrained, precompetitive consortia have a major role to play in furthering the translational aspects of informatics. As noted by Barnes *et al*¹⁰ recently, there is little competitive value in companies creating or hosting their own instances of publicly available data. In addition, storing the massive amounts of data being

produced by genome-wide association studies and other similar —omics efforts will likely become unsustainable for a single entity. With limited resources, investing in technology for technology's sake (eg, sequencing platforms) becomes less justifiable, particularly because the actual sequencing may be more efficiently completed by a third party. Although business and sustainability models remain to be fully developed, precompetitive consortia such as the Pistoia Alliance¹¹ or Alzheimer's Disease Neuroimaging Initiative¹² or government-sponsored programmes such as the EU's Innovative Medicines Initiative¹³ are successful examples of public—private partnerships. Precompetitive efforts lift the boats of all parties involved, or bring them to the same, perhaps basic, level of understanding or performance around a particular area. Then, it is up to the company itself to determine how best to build on this framework for a competitive advantage. Ultimately, the success of precompetitive efforts will be judged by their ability to produce solutions that are implemented in both industry and academic settings, with scalable and affordable support models that will facilitate their adoption.

ADOPTING STANDARDS

The call for standards in biomedical research is possibly as old as the field, and shows little sign of losing momentum. While organizations such as the Clinical Data Interchange Standards Consortium¹⁴ and Health Level 7¹⁵ have advanced clinical standards, earlier stage research remains far more muddled. Much of this stems from the rapid technological and scientific advancements requiring new standards for new entities (eg, the multiple 'minimum information' proposals);^{16–17} however, the lack of incentives for using or adopting standards, which are often unwieldy or complex, also plays an important role. Standards are often most important for data reuse, rather than for the primary purpose for which data are collected, so the benefit of their adoption often goes to parties other than the originators of the data. Making the case for researchers to change their ways is extremely difficult if making data standards compliant is seen merely as an extra step or as requiring a high level of effort for a low level of reward.^{18–19}

The current state of biomedical standards makes the case for TI both more difficult and more urgent. Systems that do not talk to each other hinder the achievement of scale that is the basis of a well-formed informatics strategy. Standards will continue to be important as complex disease models evolve involving data from multiple stages of research, multiple platforms and multiple authors. In this case, the role of TI is either the top-down approach of providing a 'universal thesaurus' that can translate among the different data sources, or the bottom-up approach of developing and promulgating standards that will be widely adopted and 'enforced' by editors, grantmakers, or regulators.

THE ROLE OF HEALTHCARE INFORMATION TECHNOLOGY

Informatics platforms also have a major role to play in the dissemination of new research results, such as findings related to comparative effectiveness, and their eventual implementation in clinical practice. To achieve high quality patient care outcomes, the constant flow of new guidelines and other new evidence-based recommendations requires sophisticated and timely decision support.^{20–22} Electronic health records (EHR) are currently the best, dynamic vehicle for bringing these findings to clinicians in a meaningful way.^{20–21} In addition to being relatively easy to update with new findings, EHR also can act as *de facto*

registries, capturing longitudinal data while following the actual impact of new recommendations that are put into practice. Data from EHR and innovative analytics can also contribute to post-marketing safety surveillance.²³

Working with integrated delivery networks, academic medical centers, or EHR vendors, the industry can contribute to design specifications that make collecting the 'right' data more easy through an EHR or other electronic system. Data capture and analysis, particularly when those data conform to accepted standards and can be shared both internally and with external partners (eg, contract research organizations) and regulators, is at the core of executing successful clinical trials. EHR or other clinical databases may also play an important role in designing or improving recruiting for clinical trials and protocols, and potentially also for data capture in clinical trials.^{24–27}

Balancing the priorities of clinical care and research is a challenge, but is also an important, concrete example of TI, in which system and database design to collect data important for research (eg, more detailed disease characterization in selected patients, in addition to the clinical data collected at a visit) can allow for easy integration into workflows and ultimately the collection of data that are beneficial for all involved.⁹

INVESTING IN INFORMATICS

Industry groups have created several translational tools that facilitate drug discovery and research. Groups within Pfizer's Computational Sciences Center of Excellence, for example, have created tools such as Targetpedia, which combines detailed genetic data with competitive intelligence, Toxicomatrix (toxicology), and the Target Opportunity Universe (chemical and biological data). transSMART is a data warehouse focusing on translational medicine developed by Johnson & Johnson,²⁸ and several pharmaceutical companies contributed to the Observational Medical Outcomes Partnership,²⁹ in cooperation with the Foundation for the National Institutes of Health.

Industry plays a vital role in taking the best of academic (and internal) discoveries, and proving their value by turning them into products that help patients. There is a great opportunity to further the field, both via precompetitive consortia and within individual companies, if informatics is seen as not only a service for computational biologists or outcomes researchers, but also as an area of investment for innovation. From a long-term perspective, industry benefits from investing in 'basic' informatics research to tackle issues around adverse event detection, safety reporting, or improved preclinical modeling, to name a few. Informatics tools can also help break down silos within institutions, such as providing shared bioinformatics infrastructure but also providing customizable tools that are fit for purpose, supporting different analytical approaches (eg, oncology vs neuroscience).

Strategies to determine what is 'important' in huge datasets will be important for advancing drug development, particularly as 'big data' become the norm. As more data are generated in a variety of settings, innovative ways to inform researchers what datasets are even available, both internally and externally, while promoting data reuse and rewarding the sharing of data,^{29–31} will be essential tools for translational research. Informatics departments can also contribute to the development of tools with intuitive user interfaces that can be effectively deployed without expert knowledge. Finally, the informatics community can continue its work around analyzing and identifying pathways and targets, particularly on predicting successful transitions from preclinical to clinical trials. The

combination of finding targets and predicting human success would also be at the core of a precision medicine strategy.

The major goal of TI is facilitating scientific research and advancement; however, another important aspect of the field is the quality of the tools and systems produced. Translation is a long, complicated process, and many tools focus on a particular stage. Both the academic and industry informatics community could benefit from looking at tools or systems they are producing and consider how they meet stakeholders' needs, particularly for those with non-technical backgrounds. In particular, is there more than academic benefit from creating a new system or promulgating a new standard instead of refining tools that already exist, making them more usable or more widely applicable? Peer reviewers, granting agencies, and tenure committees should also take the translational aspects of proposals or research into account when making decisions.

CONCLUSIONS

Several concrete, although not necessarily simple, efforts could move the field toward achieving the promise of TI. Early stage research could be facilitated by more precompetitive consortia, which would focus on bringing data sources together and opening them up for interrogation, eliminating much of the redundant work in member organizations. Data from preclinical assays to postmarketing safety surveillance could be captured using a common, well-defined, well-enforced standard across both industry and academia. Widespread adoption of healthcare information technology could allow for the rapid implementation of research findings and their evaluation in practice. Finally, as data generation across biological scales becomes easier and cheaper, informatics can provide models of increasingly complex biological systems and the effects of their perturbations, and ultimately understanding of the clinical impact of the changes to systems. These are not new recommendations, but in the current resource-constrained environment, their adoption is increasingly urgent.

Acknowledgments The author would like to thank Christoph Brockel, PhD for his comments on earlier drafts of this work.

Competing interests The author is a full-time employee of Pfizer, Inc.

Provenance and peer review Commissioned; externally peer reviewed.

REFERENCES

1. **Reis SE**, Berglund L, Bernard GR, *et al*. Reengineering the national clinical and translational research enterprise: the strategic plan of the National Clinical and Translational Science Awards Consortium. *Acad Med* 2010;**85**:463–9.
2. **Bunnage ME**. Getting pharmaceutical R&D back on target. *Nat Chem Biol* 2011;**7**:335–9.
3. **Buchan NS**, Rajpal DK, Webster Y, *et al*. The role of translational bioinformatics in drug discovery. *Drug Discov Today* 2011;**16**:426–34.
4. **Payne PR**, Embi PJ, Niland J. Foundational biomedical informatics research in the clinical and translational science era: a call to action. *J Am Med Inform Assoc* 2010;**17**:615–16.
5. **Payne PR**, Embi PJ, Sen CK. Translational informatics: enabling high-throughput research paradigms. *Physiol Genomics* 2009;**39**:131–40.
6. **Sarkar IN**, Butte AJ, Lussier YA, *et al*. Translational bioinformatics: linking knowledge across biological and clinical realms. *J Am Med Inform Assoc* 2011;**18**:354–7.
7. **Downing GJ**, Boyle SN, Brinner KM, *et al*. Information management to enable personalized medicine: stakeholder roles in building clinical decision support. *BMC Med Inform Decis Mak* 2009;**9**:44.
8. **Kawamoto K**, Lobach DF, Willard HF, *et al*. A national clinical decision support infrastructure to enable the widespread and consistent practice of genomic and personalized medicine. *BMC Med Inform Decis Mak* 2009;**9**:17.
9. **Nadler JJ**, Downing GJ. Liberating health data for clinical research applications. *Sci Transl Med* 2010;**2**:18cm6.
10. **Barnes MR**, Harland L, Foord SM, *et al*. Lowering industry firewalls: pre-competitive informatics initiatives in drug discovery. *Nat Rev Drug Discov* 2009;**8**:701–8.
11. **Pistoia Alliance**. <http://pistoiaalliance.org/industrychallenges/scientific-collaboration.html> (accessed 28 Dec 2011).
12. **Weiner MW**, Aisen PS, Jack CR Jr, *et al*. The Alzheimer's disease neuroimaging initiative: progress report and future plans. *Alzheimers Dement* 2010;**6**:202–11 e7.
13. **Innovative Medicines Initiative**. <http://www.imi.europa.eu/> (accessed 28 Dec 2011).
14. **Clinical Data Interchange Standards Consortium**. <http://www.cdisc.org/> (accessed 28 Dec 2011).
15. **Health Level Seven International**. <http://www.hl7.org/> (accessed 28 Dec 2011).
16. **Orchard S**, Al-Lazikani B, Bryant S, *et al*. Minimum information about a bioactive entity (MIBBE). *Nat Rev Drug Discov* 2011;**10**:661–9.
17. **Orchard S**, Taylor CF. Debunking minimum information myths: one hat need not fit all. *N Biotechnol* 2009;**25**:171–2.
18. **Nelson B**. Data sharing: empty archives. *Nature* 2009;**461**:160–3.
19. **Santos C**, Blake J, States DJ. Supplementary data need to be kept in public repositories. *Nature* 2005;**438**:738.
20. **Cebul RD**, Love TE, Jain AK, *et al*. Electronic health records and quality of diabetes care. *N Engl J Med* 2011;**365**:825–33.
21. **Ebell M**. AHRQ White Paper: use of clinical decision rules for point-of-care decision support. *Med Decis Making* 2010;**30**:712–21.
22. **Lyman JA**, Cohn WF, Bloomrosen M, *et al*. Clinical decision support: progress and opportunities. *J Am Med Inform Assoc* 2010;**17**:487–92.
23. **Linder JA**, Haas JS, Iyer A, *et al*. Secondary use of electronic health record data: spontaneous triggered adverse drug event reporting. *Pharmacoepidemiol Drug Saf* 2010;**19**:1211–15.
24. **El Fadly A**, Rance B, Lucas N, *et al*. Integrating clinical research with the healthcare enterprise: from the RE-USE project to the EHR4CR platform. *J Biomed Inform* 2011;**44**(Suppl 1):S94–S102.
25. **Heinemann S**, Thuring S, Wedeken S, *et al*. A clinical trial alert tool to recruit large patient samples and assess selection bias in general practice research. *BMC Med Res Methodol* 2011;**11**:16.
26. **McCarty CA**, Chisholm RL, Chute CG, *et al*. The eMERGE Network: a consortium of biorepositories linked to electronic medical records data for conducting genomic studies. *BMC Med Genomics* 2011;**4**:13.
27. **Ohmann C**, Kuchinke W. Future developments of medical informatics from the viewpoint of networked clinical research. Interoperability and integration. *Methods Inf Med* 2009;**48**:45–54.
28. **Szalma S**, Koka V, Khasanova T, *et al*. Effective knowledge management in translational medicine. *J Transl Med* 2010;**8**:68.
29. **Stang PE**, Ryan PB, Racoosin JA, *et al*. Advancing the science for active surveillance: rationale and design for the Observational Medical Outcomes Partnership. *Ann Intern Med* 2010;**153**:600–6.
30. **Abernethy DR**, Woodcock J, Lesko LJ. Pharmacological mechanism-based drug safety assessment and prediction. *Clin Pharmacol Ther* 2011;**89**:793–7.
31. **Botsis T**, Hartvigsen G, Chen F, *et al*. Secondary use of EHR: data quality issues and informatics opportunities. *AMIA Summits Transl Sci Proc* 2010;**1**:1–5.