BIOMEDICAL INFORMATICS AND TRANSLATIONAL MEDICINE

Ramshankar Varma

Assistant Teacher, Shri S K Somaiya Vinay Mandir and Jr. College of Arts, Science and Commerce, Vidyavihar, Mumbai 400077, India

ABSTRACT

Biomedical informatics involves a core set of methodologies that can provide a foundation for crossing the "translational barriers" associated with translational medicine. To this end, the fundamental aspects of biomedical informatics (e.g., bioinformatics, imaging informatics, clinical informatics, and public health informatics) may be essential in helping improve the ability to bring basic research findings to the bedside, evaluate the efficacy of interventions across communities, and enable the assessment of the eventual impact of translational medicine innovations on health policies. Here, a brief description is provided for a selection of key biomedical informatics topics (Decision Support, Natural Language Processing, Standards, Information Retrieval, and Electronic Health Records) and their relevance to translational medicine. Based on contributions and advancements in each of these topic areas, the article proposes that biomedical informatics practitioners ("biomedical informaticians") can be essential members of translational medicine teams.

Keywords: Internet utility, Medical diagnosis, biomedical information technology

INTRODUCTION

Biomedical informatics, by definition, incorporates a core set of methodologies that are applicable for managing data, information, and knowledge across the translational medicine continuum, from bench biology to clinical care and research to public health. To this end, biomedical informatics encompasses a wide range of domain specific methodologies. In the present discourse, the specific aspects of biomedical informatics that are of direct relevance to translational medicine are: (1) bioinformatics; (2) imaging informatics; (3) clinical informatics; and, (4) public health informatics. These support the transfer and integration of knowledge across the major realms of translational medicine, from molecules to populations. A partnership between biomedical informatics and translational medicine promises the betterment of patient carethrough development of new and better understood interventions used effectively in clinics as well as development of more informed policies and clinical guidelines. The ultimate goal of translational medicine is the development of new treatments and insights towards the improvement of health across populations. The first step in this process is the identification of what interventions might be worthy to consider. Next, directed evaluations (e.g., randomized controlled trials) are used to identify the efficacy of the intervention and to provide further insights into why a proposed intervention works. Finally, the ultimate success of an intervention is the identification of how it can be appropriately scaled and applied to an entire population. The various contexts presented across the translational medicine spectrum enable a "grounding" of biomedical informatics approaches by providing specific scenarios where knowledge management and integration approaches are needed. Between each of these steps, translational barriers are comprised of the challenges associated with the translation of innovations developed through bench-based experiments to their clinical validation in bedside clinical trials, ultimately leading to their adoption by communities and potentially leading to the establishment of policies. It is particularly important to emphasize that, while the major thrust is in the forward direction, accomplishments, and setbacks can be used to valuably inform both sides of each translational barrier.

An important enabling step to cross the translational barriers is the development of trans-disciplinary teams that are able to integrate relevant findings towards the identification of potential breakthroughs in research and clinical intervention. To this end, biomedical informatics professionals ("biomedical informaticians") may be an essential addition to a translational medicine team to enable effective translation of concepts between team members with heterogeneous areas of expertise. Translational medicine teams will need to address many of the challenges that have been the focus of biomedical informatics since the inception of the field. What follows is a brief description of biomedical informatics, followed by a discussion of selected key topics that are of relevance for translational medicine: (1) Decision Support; (2) Natural Language Processing; (3) Standards; (4) Information Retrieval; and, (5) Electronic Health Records. For each topic, progress and activities in bio-, imaging, clinical and public health informatics are described. The article then concludes with a consideration of the role of biomedical informaticians in translational medicine teams.

BIOMEDICAL INFORMATICS

Biomedical informatics is an over-arching discipline that includes sub-disciplines such as bioinformatics, imaging informatics, clinical informatics, and public health informatics; the relationships between the sub-disciplines have been previously well characterized, and are still tenable in the context of translational medicine. Much of the identified synergy between biomedical informatics and translational medicine can be organized into two major categories that build upon the sub-disciplines of biomedical informatics: (1) translational bioinformatics (which primarily consists of biomedical informatics methodologies aimed at crossing the T1 translational barrier) and (2) clinical research informatics (which predominantly consists of biomedical informatics techniques from the T1 translational barrier across the T2 and T3 barriers). It is important to emphasize that the role of biomedical informatics in the context of translational medicine is not to necessarily create "new" informatics techniques. Instead, it is to apply and advance the rich cadre of biomedical informatics approaches within the context of the fundamental goal of translational medicine: facilitate the application of basic research discoveries towards the betterment of human health or treatment of disease.

Clinical informatics has historically been described as a field that meets two related, but distinct needs: patient-centric and knowledge-centric. This notion can be generalized for all of biomedical informatics within the context of translational medicine to suggest that the goals are either to meet the needs of user-centric stakeholders (e.g., biologists, clinicians, epidemiologists, and health services researchers) or knowledge-centric stakeholders (e.g., researchers or practitioners at the bench, bedside, community, and population level). Bioinformatics approaches are needed to identify molecular and cellular regions that can be targeted with specific clinical interventions or studied to provide better insights to the molecular and cellular basis of disease. Imaging informatics techniques are needed for the development and analysis of visualization approaches for understanding pathogenesis and identification of putative treatments from the molecular, cellular, tissue or organ level. Clinical informatics innovations are needed to improve patient care through the availability and integration of relevant information at the point of care. Finally, public health informatics solutions are required to meet population based needs, whether focused on the tracking of emergent infectious diseases, the development of resources to relate complex clinical topics to the general population or the assessment of how the latest clinical interventions are impacting the overall health of a given population.

At the T1 translational barrier crossing, translational bioinformatics is rapidly evolving with the enhancement and specialization of existing bioinformatics techniques and biological databases to enable identification of specific bench-based insights. Similarly, clinical research informatics emphasizes the use of biomedical informatics approaches to enable the assessment and moving of basic science innovations from the T1 translational barrier and across the T2 and T3 translational barriers (as depicted in Figure 1). These approaches may involve the enhancement and specialization of existing and new clinical and public health informatics techniques within the context of implementation and controlled assessment of novel interventions, development of practice guidelines, and outcomes assessment.

TRANSLATIONAL BIOINFORMATICS

The future of biomedical informatics depends on the ability to leverage common frameworks that enable the translation of research hypotheses into practical and proven treatments. Progress has already been seen in the development of knowledge management infrastructures and standards to enable biomedical research to facilitate general research inquiry in specific domains (e.g., cancer and neuroimaging]). It is also imperative for such advancements to be done in the context of improving user-centric needs, thereby improving patient care. To this end, the ability to manage and enable exploration of information associated with the biomedical research enterprise suggests that human medicine may be considered as the ultimate model organism. Towards this aspiration, biomedical informaticians are uniquely equipped to facilitate the necessary communication and translation of concepts between members of trans-disciplinary translational medicine teams.

NATURAL LANGUAGE

Processing Natural Language Processing (NLP) systems fall into two general categories: (1) natural language understanding systems that extract information or knowledge from human language forms (either text or speech), often resulting in encoded and structured forms that can be incorporated into subsequent applications; and, (2) natural language generation systems that generate human understandable language from machine representations (e.g., from within a knowledge bases or systems of logical rules). NLP has a strong relationship to the field of computational linguistics, which derives computational models for phenomena associated with natural language (encapsulated as either sets of handcrafted rules or statistically derived models). The development and application of NLP approaches has been a significant focus of research across the entire spectrum of biomedical informatics. Biological knowledge extraction has also been a major area of focus in NLP systems, including the use of NLP methods to facilitate the prediction of molecular pathways. Within imaging informatics, there has been a range of applications that involve processing and generating information associated with clinical images that are often used to help summarize and organize radiology images. In clinical informatics, there have been great advances in the extraction of information from semi-structured or unstructured narratives associated with patient care as well as the development of applications for generating summaries or reports automatically. In the realm of public health, NLP approaches have been demonstrated to facilitate the encoding and summarization of significant information at the population level, such as for describing functional status and outbreak detection. Peer-reviewed literature, such as indexed by MEDLINE, has been shown to be a source of previously unknown inferences across domains as well as linkages between the bioinformatics and clinical informatics communities. In addition to MEDLINE, which grows by approximately 1 million citations per year, the increasing adoption of Electronic Health Records will lead to increased volumes of natural language text.

To this end, NLP approaches will increasingly be needed to wade through and systematically extract and summarize the growing volumes of textual data that will be generated across the entire translational spectrum. There has also been some work in NLP that directly strives to develop linkages across disparate text sources (e.g., bridging e-mail communications to relevant literature). Within the realm of translational medicine, NLP approaches will be increasingly poised to facilitate the development of linkages between unstructured and structured knowledge sources across the realms of biology, medicine, and public health.

Standards that have been developed and are implemented by the biomedical informatics community will be an essential component towards the goal of integrating relevant data across the translational barriers (e.g., to answer questions like what is the comparative effectiveness of a particular pharmacogenetic treatment versus conventional pharmaceutical treatments in the general population?). Additionally, standards can facilitate the access and integration of information associated with a particular individual in light of available biological, imaging, clinical, and public health data (including improved access to these data from within medical records), ultimately enabling the development and testing the utility of "personalized medicine." Consequently, translational medicine will depend on biomedical informatics approaches to leverage existing standards (e.g., MIAME, HL7, and DICOM) and resources like the UMLS, in addition to developing new standards for specialized domains (e.g., cancer and neuroimaging)

INFORMATION RETRIEVAL

Information retrieval systems are designed for the organization and retrieval of relevant information from databases. The basic premise is that a query is presented to a system that then attempts to retrieve the most relevant items from within database(s) that satisfy the request. The quality of the results is then measured using statistics such as precision (the number of relevant results retrieved relative to the total number of retrieved results) and recall (the number of relevant results retrieved relative to the total number of relevant items in the database). Across the field of biomedical informatics, various efforts have focused on the need to bring together information across a range of data sources to enable information retrieval queries. Perhaps the most popular information retrieval tool is the PubMed interface to the MEDLINE citation database that contains information across much of biomedicine. In addition to MEDLINE, the growth of publicly available resources has been especially remarkable in bioinformatics, which generally focus on the retrieval of relevant biological data (e.g., molecular sequences from GenBank given a nucleotide or protein sequence). Information retrieval systems have also been developed in bioinformatics that are able to retrieve relevant data from across multiple resources simultaneously (e.g., for generating putative annotations for unknown gene sequences. Imaging information retrieval systems have been a rich research area where relevant images are retrieved based on image similarity (e.g., to identify pathological images that might be related to a particular anatomical shape and related clinical context). Within clinical environments, information retrieval systems have been developed that can link users to relevant clinical reference resources based on using the particular clinical context as part of the query (e.g., to identify relevant articles based on a specific abnormal laboratory result). Information retrieval systems have been developed in public health to identify relevant information for consumers, epidemiologists, and health service researchers given varying types of queries. The procedural tasks involved with information retrieval often involve natural language processing and knowledge representation techniques, such as highlighted previously.

The integration of natural language processing, knowledge representation and information retrieval systems has led to the development of "question-answer" systems that have the potential to provide more user-friendly interfaces to information retrieval system.

ELECTRONIC HEALTH RECORDS

Medical charts contain the sum of information associated with an individual's encounters with the health care system. In addition to data recorded by direct care providers (e.g., physicians and nurses), medical charts typically include data from ancillary services such as radiology, laboratory, and pharmacy. With the increasing electronic availability of data across the health care enterprise, paper-based medical charts have evolved to become computerized as Electronic Health Records (EHRs). EHRs can capture a variety of information (e.g., by clinicians at the bedside) and have electronic interfaces to individual services (e.g., administrative, laboratory, radiology, and pharmacy). Many EHRs can enable Computerized Provider Order Entry (CPOE), which allows clinicians to electronically order services and may also enable real-time clinical decision support (e.g., provide an alert about an order that could lead to an adverse event). Clinical documentation can be entered directly into EHR systems, allowing for potentially fewer issues due to transcription delays or difficulty in deciphering handwritten notes. An artifact of EHRs is the development of more robust clinical and research data warehouses, which can be used for subsequent studies.

The increased availability of electronic health data, which are largely available and organized within EHRs, may have a significant impact on translational medicine. For example, the emergence of "personal health" projects (e.g., Google Health]) and consumer services (e.g., 23andMe) has the potential to generate more genotype (i.e., "bench") and phenotype (i.e., "bedside") data that may be analyzed relative to community-based studies. The raw elements that could lead to the next breakthroughs may be made available as part of the data deluge associated with consumer-driven, "grass-roots" efforts. Such initiatives, in addition to the other core biomedical informatics topics discussed here (decision support, natural language processing, and information retrieval techniques), will enable the leveraging of EHRbased health data to catalyze the crossing of the translational barriers.

THE ROLE OF THE BIOMEDICAL INFORMATICIAN IN A TRANSLATIONAL MEDICINE TEAM

Translational medicine is a trans-disciplinary endeavor that aims to accelerate the process of bringing innovations into practice through the linking of practitioners and researchers across the spectrum of biomedicine. As evidenced by major funding initiatives (e.g., the United States National Institutes of Health "Roadmap"), there is great hope in the development of a new paradigm of research that catalyzes the process from bench to practice. The trans-disciplinary nature of the translational barrier crossings in translational medicine endeavors will increasingly necessitate biomedical informatics approaches to manage, organize, and integrate heterogeneous data to inform decisions from bench to bedside to community to policy. The distinctions between multi-disciplinary, inter-disciplinary, and trans-disciplinary goals have been described as the difference between additive, interactive, and holistic approaches. Unlike multi-disciplinary or inter-disciplinary endeavors, trans-disciplinary initiatives must be completely convergent towards the development of completely new research paradigms. The greatest challenge faced by translational medicine, therefore, is the difficulty in truly being a trans-disciplinary science that brings together researchers and practitioners that traditionally work within their own "silos" of practice.

Formally trained biomedical informaticians have a unique education, often with domain expertise in at least one area, which is specifically designed to enable trans-disciplinary team science, such as needed for the success within a translational medicine team. There is some discussion over what level of training constitutes the minimal requirements for biomedical informatics training, including discussion about what combination of technical and non-technical skills are needed. Their ability to interact with key stakeholders across the translational medicine spectrum (e.g., biologists, clinicians/clinical researchers, epidemiologists, and health services researchers). Furthermore, biomedical informaticians bring the methodological approaches (depicted as the shadowed region in Figure 2), such as the five topics highlighted in earlier sections of this article, which can enable the development and testing of new trans-disciplinary hypotheses. It is important to note that the topics discussed in this article are only a sampling of the full array of biomedical informatics techniques that are available (e.g., cognitive science approaches, systems design and engineering, and telehealth). The success of translational medicine will depend not only on the addition of biomedical informaticians to translational medicine teams, but also on the acceptance and understanding of what biomedical informatics consists of by other members in the team. To this end, the importance of biomedical informatics training has been underscored as a key area of required competency across the spectrum of translational medicine, from biologists to clinicians to public health professionals. There has been some demonstrable success in the development of experiences that focus on training "agents of change" with necessary core concepts as well as hallmark distributed educational programs that aim to provide formal educational opportunities for biomedical informatics training. The composition of translational medicine teams will also depend on the appropriate intermixing of biomedical informatics expertise to complement the requisite domain expertise. To this end, the success of translational medicine endeavors may undoubtedly be greatly enhanced with biomedical informatics approaches; however, the appropriate synergistic relationship between biomedical informaticians and other members of the translational medicine team remains one of the next major challenges to be addressed in pursuit of translational medicine breakthroughs.

CONCLUSION

Since its beginnings, biomedical informatics innovations have been developed to support the needs of various stakeholders including biologists, clinicians/clinical researchers, epidemiologists, and health services researchers. A range of biomedical informatics topics, such as those described in this paper, form a suite of elements that can transform data across the translational medicine spectrum. The inclusion of biomedical informaticians in the translational medicine team may thus help enable a trans-disciplinary paradigm shift towards the development of the next generation of ground breaking therapies and interventions.

REFERENCES

1. Shortliffe EH, Cimino JJ: Biomedical informatics: computer applications in health care and biomedicine New York, NY: Springer, 3 2006.
2. Greenes RA, Shortliffe EH: Commentary: Informatics in biomedicine and health care. Acad Med 2009, 84:818-820.
3. Bernstam EV, Hersh WR, Johnson SB, Chute CG, Nguyen H, Sim I, Nahm M, Weiner MG, Miller P, DiLaura RP, Overcash M, Lehmann HP, Eichmann D, Athey BD, Scheuermann RH, Anderson N, Starren J, Harris PA, Smith JW, Barbour E, Silverstein JC, Krusch DA, Nagarajan R, Becich MJ: Synergies and distinctions between computational disciplines in biomedical research: perspective from the Clinical andTranslational Science Award programs. Acad Med 2009, 84:964-970.
4. Collen MF: The origins of informatics. J Am Med Inform Assoc 1994, 1:91-107.
5. Collen MF: Fifty years in medical informatics. Yearb Med Inform 2006, 174-179.
6. Haux R: Individualization, globalization and health–about sustainable information technologies and the aim of medical informatics. Int J Med Inform 2006, 75:795-808. 7. Kuhn KA, Knoll A, Mewes HW, Schwaiger M, Bode A, Broy M, Daniel H, Informatics and medicine–from molecules to populations. Methods Inf Med 2008, 47:283-295.
7. Embi PJ, Kaufman SE, Payne PR: Biomedical informatics and outcomes research: enabling knowledge-driven health care. Circulation 2009, 120:2393-2399.
8. Payne PR, Johnson SB, Starren JB, Tilson HH, Dowdy D: Breaking the translational barriers: the value of integrating biomedical informatics and translational research.
9. J Investig Med 2005, 53:192-200. 10. Payne PR, Embi PJ, Sen CK: Translational informatics: enabling highthroughput research paradigms. Physiol Genomics 2009, 39:131-140.
10. Woolf SH: The meaning of translational research and why it matters. JAMA 2008, 299:211-213.