Introduction to Clinical Research Informatics

1

Kate Fultz Hollis, Rachel L. Richesson, and James E. Andrews

Abstract

This chapter introduces in this third edition of *Clinical Research Informatics* the overview of important constructs and methods within the subdomain of clinical research informatics now. The chapter sets the tone and scope for the text, highlights important topics and themes, and describes the content and organization of chapters.

Keywords

Clinical research informatics definition · CRI · Theorem of informatics · American Medical Informatics Association · Biomedical informatics

K. F. Hollis (✉)
Department of Medical Informatics and Clinical Epidemiology, Oregon Health & Science University, Portland, OR, USA
e-mail: fultzhol@ohsu.edu

R. L. Richesson
Department of Learning Health Sciences, University of Michigan Medical School, Ann Arbor, MI, USA
e-mail: richessr@med.umich.edu

J. E. Andrews
School of Information, College of Arts and Sciences, University of South Florida, Tampa, FL, USA
e-mail: jimandrews@usf.edu

Overview

Welcome to the third edition of *Clinical Research Informatics*. It has been 10 years since co-editors Rachel Richesson and James Andrews published the first edition of this book, and while informatics foundations in clinical research are roughly the same, the tremendous change in technology brings us more data to use, advanced methods in data analysis, and increased public awareness and appreciation of the growing roles of technology in medicine and research. We have also seen increased public awareness of the importance of biomedical research and growing appreciation of the challenges and limitations of our current national research infrastructure. As we have done in past editions, we begin with definitions.

Clinical research is the branch of medical science that investigates the safety and effectiveness of medications, devices, diagnostic products, and treatment regimens intended for human use in the prevention, diagnosis, treatment, or management of a disease. The documentation, representation, and exchange of information in clinical research are inherent to the very notion of research as a controlled and reproducible set of methods for scientific inquiry. Contemporary clinical research actually represents new application of statistics to medicine with the acceptance of randomized controlled clinical trials as the gold standard [1] only recognized in this last 70 years. Clinical research has been characterized as a discipline

R. L. Richesson et al. (eds.), *Clinical Research Informatics*, Health Informatics,
https://doi.org/10.1007/978-3-031-27173-1_1

resting on three pillars of principle and practice related to control, mensuration, and analysis [2], though these can be more modernly interpreted as a triad of expertise in medicine, statistics, and logistics [3].

Clinical research informatics (CRI) is the application of informatics principles and technologies to support the spectrum of activities and business processes that represent clinical research. Informatics, broadly defined as the intersection of information and computer science with a health-related discipline, has a foundation that has drawn from many well-established, theory-based disciplines, including computer science, library and information science, cognitive science, psychology, and sociology. The fundamental theorem of informatics [4] states that humans plus information technology should function and perform better together than humans alone, and so informatics is a source for supportive technologies and tools that enhance—but not replace—unreservedly human processes.

The US National Institutes of Health (NIH) offers a comprehensive and widely accepted definition for clinical research that includes a spectrum of populations, objectives, methods, and activities. Specifically, this broad definition states that "clinical research is… patient-oriented research conducted with human subjects (or on material of human origin that can be linked to an individual)" [5]. Under this definition, clinical research includes investigation of the mechanisms of human disease, therapeutic interventions, clinical trials, development of new technologies, epidemiology, behavioral studies, and outcomes and health services research.

The challenges in clinical research—and, thus, the opportunities for informatics support—arise from many different objectives and requirements, including the need for optimal protocol design, regulatory compliance, improved patient recruitment and engagement, efficient protocol management, and data collection and acquisition; data storage, transfer, processing, and analysis; and, impeccable patient safety throughout. Regardless of clinical domain or study design, high-quality data collection and standard, formalized data representation are critical to the fundamental notion of research reproducibility.

The need for clinical research informatics is stronger and more visible that ever before as we are emerging from a global pandemic caused by a novel and lethal virus. As the *New England Journal of Medicine* reported, in December 2019, a cluster of patients with pneumonia of unknown cause was linked to a seafood wholesale market in Wuhan, China. A previously unknown betacoronavirus was discovered through the use of unbiased sequencing in samples from patients with pneumonia [6]. Since that time, the virus identified as COVID-19 has claimed millions of lives, and the pandemic opened greater awareness of research, as there was a global push to develop and evaluate diagnostic tests, vaccines, and ultimately treatments. The gravity of the pandemic not only made progress faster than usual (the expeditious testing of the vaccine efficacy was unprecedented [7]) but also cast a bright light on the many challenges of clinical research and significant limitations of our infrastructure. The experience has demonstrated the essential role of clinical research informatics to support faster research in real-world settings as focus on pandemic vaccines and treatments was intense. In parallel, the FDA, in recognition that the current clinical research infrastructure and approach are not sufficient to address and inform the growing number of clinical questions and emerging conditions/therapeutics (and also recognizing the need for real-world evidence (RWE) to inform real-world practice), has called for the development of new methods for using real-world data (RWD) [8].

Clinical research investigations of other conditions were also impacted by the COVID-19 pandemic. Many clinical trials had to be modified and delayed according to COVID-19 protocols (following institutional guidelines and the Federal Drug Administration's "Conduct of Clinical Trials of Medical Products During the COVID-19 Public Health Emergency" [9], but critical work did continue and has informed new methods adapting trials and analysis to cope with disruption and competing resources [10].

Contexts and Attempts to Define Clinical Research Informatics

Original research is needed to create an evidence base for information and communications technologies that meaningfully address the business needs of research and also streamline, change, and improve the business of research itself. CRI has now become established: several books—including this one—are now available, CRI-focused working groups in the American Medical Informatics Association (AMIA) and HL7 are quite active, and CRI is now a regular section in the International Medical Informatics Association (IMIA) Yearbook each year. Standards and best practices for research have started to emerge, as are standards for education and training in the field [11].

The comprehensive and definitive definition for CRI is presented by Embi and Payne (2009) who characterize CRI as "the sub-domain of biomedical informatics concerned with the development, application, and evaluation of theories, methods, and systems to optimize the design and conduct of clinical research and the analysis, interpretation, and dissemination of the information generated" [12]. A more descriptive definition is offered by AMIA, where CRI includes evaluation and modeling of clinical and translational research workflow; social and behavioral studies involving clinical research; designing optimal human–computer interaction models for clinical research applications; improving and evaluating information capture and data; flow in clinical research; optimizing research site selection, investigator, and subject recruitment; knowledge engineering and standards development as applied to clinical research; facilitating and improving research reporting to regulatory agencies; and enhancing clinical and research data mining, integration, and analysis. (See AMIA CRI Working Group: https://amia.org/community/working-groups/clinical-research-informatics.) The definition and illustrative activities of CRI both emerged from in-person and virtual meetings and interviews with self-identified CRI practitioners within the AMIA organization. Since CRI was formally defined in 2009, the number of self-identified CRI practitioners continues to grow in academic medical centers, pharmaceutical companies, and health IT and innovation start-ups. The growing demand to use EHR data to improve the efficiency, generalizability, and relevance of clincal research is driving new models of patient engagement—both in research and with their health data—resulting in tremendous growth in mobile applications, further broadening the landscape of CRI innovation and practice.

The scope and number of clinical and research questions to be addressed by CRI have obviously evolved over time and since our first two editions of this book. A single professional or educational home for CRI, and as such a source to develop a single consensus and more precise definition, is still lacking at present and likely unachievable given the multidisciplinary and multinational and multicultural scope of CRI activities. However, the AMIA CRI Working Group remains a crucial and dominate forum for CRI, providing a more detailed articulation of the role of the chief research information officer (CRIO) [13]. What is important to note is that this is all reflective of the bottom-up development of our field. For the annual survey of CRI for the IMIA Yearbook in 2020, Anthony Solomonides wrote that many topics of interest today may also have been featured 10 years ago, but some topics did not figure at all in 2009. Among those new topics noted are artificial intelligence (AI), especially in the form of machine/deep learning, our understanding of causal inference, and the shifting trend in the use of "real-world evidence" (RWE) often gathered by networks using one or other of the several "common data models" [14].

The first references to what is now known as clinical research informatics go back to the 1960s and predict the inevitable use of computers to support data collection and analysis in research [15]. Christopher Chute (Chap. 2) goes into more detail on historical changes in research in terms of volume and types of data in the physical sciences. The use of clinical databases for research inquiry was first established in the late 1960s, and by the next decade, there were at least a handful of clinical information systems being

used for research. This history is well described by Morris F. Collen in a 1990 historical review [16]. In short course, it was clear that structured data entry and data standards would be a critical component of any computerized support or analysis system in research [17]. M. Scott Blois was the first to describe the complexities and dimensions of representing medical data in his seminal book [18]. Blois first recognized that systems could and should support more than queries about a *single* patient's data but rather should be searchable to retrieve *many* patient records to support research and quality monitoring. The first applications focused on retrieval of clinical information to identify and understand patient subpopulations [19]. Others saw the potential for tapping these clinical databases in observational research and knowledge discovery; by the 1970s, cancer and tumor registries were well established, and cardiovascular disease registries emerged. For the first few decades, computers in clinical research were indeed centered around maintaining a database focused on collecting and querying clinical data. The advent of patient eligibility screening and trial recruitment systems in the 1990s represents the introduction of computers to support clinical research *processes* [20–22]. The regulated nature of human trials, especially since the formal inquiry and establishment of standards for the field in the 1970s, created a critical need for documentation of methods and processes, as well as analyses and findings, and we saw systems emerge in the late 1980s that began to address the conduct of studies. The capabilities of these systems have improved, and their use has proliferated. Now, clinical research management systems of various types support the collection of data and the coordination of research tasks. The primary functionality of commercial applications today is essentially concerned with the delivery of valid and accurate data in conformity with the Good Clinical Practice (GCP) guidelines [23], and in most cases, these systems are still not well integrated with patient care systems.

The enormity of data generated from new diagnostic and measurement technologies, the increasing ability to collect data rapidly from patients or external data sources, and the scope and scale of today's research enterprises have led to a bewildering array and amount of data and information. Information technology has contributed to the information management problems by generating more data and information, but the techniques and principles derived from informatics promise to purposively utilize IT to address the issues of data collection, information management, process and protocol management, communication, and knowledge discovery, as well as show promise to improve research efficiencies, increase our knowledge of therapeutic evaluation, and impact human health and the global economy. Still, in time, these tools will need to be evaluated via more formal means and evolve or be replaced by the next generation of tools and methods. As original informatics research and proper system evaluations—including randomized trials of various systems with outcome measures related to research efficiency, quality, and patient safety—are conducted, published, and scrutinized, *evidence-based* research systems, practices, and tools can be deployed and subsequently increase our ability to generate the knowledge and clinical evidence needed to address pressing and emergent public health problems.

Perspective, Objectives, and Scope

This book comes during a very challenging time for CRI and biomedical informatics. Since the first and second editions of this text, we have seen new legislation (*twenty-first Century Cures Act*) and new programs including the NIH's *All of Us Research Program* that show promise to leverage CRI to impact human health in unprecedented ways. We also have the National COVID Cohort Collaborative (N3C), a consortium with collaborators who contribute and use COVID-19 clinical data to answer critical research questions to address the pandemic (National Center for Advancing Translational Sciences https://ncats.nih.gov/n3c/about/program-faq). There is a growing interest around RWE for treatments and implementing—and generating—evidence in

Learning Health Systems (such as that defined by Agency for Healthcare Research and Quality): Learning Health Systems | Agency for Healthcare Research and Quality https://www.ahrq.gov/professionals/systems/learning-health-systems/index.html. And the rapid development and adoption of the HL7 FHIR standard brings unprecedented potential and opportunity for sharing and use of clinical data in research.

This collection of chapters is meant to represent the current knowledge in the field with an eye toward the future, and we have several new chapters including Biorepositories, Best Practices for Research Data Management, and Apps in Clinical Research. In this book, we offer foundational coverage of key areas, concepts, constructs, and approaches of medical informatics as applied to clinical research activities, in both current settings and in light of emerging policies, so as to serve as but one contribution to the discourse going on within the field now. We do not presume to capture the entirety of the field (can any single text truly articulate the full spectrum of a discipline?), but rather an array of both foundational and emerging areas that will impact clinical research and, so, CRI. Our aim is not to provide an introductory book on informatics, as is best done by Shortliffe, Cimino, and Chiang, in their foundational biomedical informatics text [24] or Hersh [25]. Rather, this text is targeted toward those who possess a basic understanding of the health informatics field and who would like to apply informatics principles to clinical research problems and processes. Many of these theories and principles presented in this text are, naturally, common across biomedical informatics and not unique to CRI; however, the authors have put these firmly in the context of how these apply to clinical research.

The excitement of such a dynamic area is fueled by the significant challenges the field must face. At this stage, there is still no consistent or formal reference model (e.g., curriculum models supporting graduate programs or professional certification) that represents the core knowledge and guides inquiry. Several informatics graduate programs across the country offer courses in clinical research informatics (Oregon Health and Science University, Columbia University, Duke University to name a few) and anecdotal reports from educators who have told us that this is the text they use. In fact, the impetus for creating *Clinical Research Informatics* came, in part, from requests from instructors for a text that offers students and others a range of knowledge and best practices from some of the top CRI scholars and practitioners. As the discipline of CRI grows, this book continues to stand out as the primary, authoritative text on the field. In this text, we attempt to cover the range of knowledge topics and practices for relevant to CRI as well as to identify several broad themes that undoubtedly will influence the future evolution of CRI.

In compiling works for this book, we were well aware that our selection of topics and placement of authors, while not arbitrary, was inevitably subjective. Others in CRI might or might not agree with our conceptualization of the discipline. Our goal is not to restrict CRI to the framework presented here; rather, that this book will stir a discourse as this subdiscipline continues to evolve. In a very loose sense, this text represents a bottom-up approach to organizing this field. There is not one exclusive professional venue for clinical research informatics; therefore, no one single place to scan for relevant topics. Numerous audiences, researchers, and stakeholders have emerged from the clinical research side (professional practice organizations, academic medical centers, the FDA and NIH sponsors, research societies like the Society for Clinical Trials, and various clinical research professional and accrediting organizations such as the Association of Clinical Research Professionals) and also from the informatics side (AMIA). Every year since 2011, Dr. Peter Embi conducts a systematic review of innovation and science of CRI and presents it to AMIA [26].

The authors for each of the chapters were selected for their demonstrated expertise in the field. We asked authors to attempt to address multiple perspectives, to paint major issues, and, when possible, to include international perspectives. Each of the outstanding authors succeeded, in our opinion, in presenting an overview of principles, objectives, methods, challenges, and

issues that currently define the topic area and that are expected to persist over the next decade. The individual voice of each author distinguishes one chapter from the other. Although some topics can be quite discreet, others overlap significantly at certain levels, and therefore, some topcis are included across multiple chapters.

Organization of the Book

This new edition has evolved over the previous editions and has been significantly reorganized. New sections have been created (increased from three to five) to offer a more logical organization expansion of related topics and chapters.

We organize the chapters under unifying themes at a high level using five broad sections: (1) the foundations of clinical research informatics; (2) enabling frameworks, processes and tools; (3) managing different types of data across clinical and translational research; (4) knowledge representation and data-driven discovery in CRI, a section that represents the future of clinical research, health, and clinical research informatics; and (5) evolving models and new opportunities for the transformation of clinical research. A new feature of this edition is the addition of "Learning Objectives" at the beginning of each chapter that highlight key concepts to help guide readers.

Part I: Foundations of Clinical Research Informatics

Chapter 2: From Notations to Data: The Digital Transformation of Clinical Research

Chapter 3: Methodological Foundations of Clinical Research

Chapter 4: The Clinical Research Environment

Chapter 5: Next Generation Biorepository Informatics: Supporting Genomics, Imaging, and Innovations in Spatial Biology

Chapter 6: Study Protocol Representation

Chapter 7: Clinical Research Information Systems

Chapter 8: Public Policy Issues in Clinical Research Informatics

The first section addresses the historical context, settings, wide-ranging objectives, and basic definitions and topics for clinical research informatics. In this section, we sought to introduce the context of clinical research and the relevant pieces of informatics that together constitute the *space* for applications, processes, problems, issues, etc. that collectively comprise CRI activities. We start with a historical perspective from Christopher Chute, whose years of experience in this domain and informatics, generally, allow for an overview of the evolution from notation to digitization. His chapter brings in historical perspectives to the evolution and changing paradigms of scientific research in general and specifically on the ongoing development of clinical research informatics. Also, the business aspects of clinical research are described and juxtaposed with the evolution of other scientific disciplines, as new technological advances greatly expanded the availability of data in those areas. Chute also illustrates the changing sociopolitical and funding atmospheres and highlights the dynamic issues that will impact the definition and scope of CRI moving forward. Extending the workflow and information needs is an overview of study designs in the chapter on methodological foundations of clinical research presented by Antonella Bacchieri and Giovanni Della Cioppa. They provide a broad survey of various research study designs (which was described in much more detail in a separate Springer text [27] authored by them) and highlight the data capture and informatics implications of each. Philip Payne follows the method chapter with a chapter focused on the complex nature of clinical research workflows included in the clinical research environment—including a discussion on stakeholder roles and business activities that make up the field. This is a foundational chapter as it describes the people and tasks that information and communication technologies (informatics) are intended to support.

At the crux of clinical research informatics is a variety of information management systems, which are characterized and described by Prakash Nadkarni. His chapter also gives a broad overview of system selection and evaluation issues. In addition, this chapter provides brief descriptions of each group of activities, system requirements

for each area, and the type and status of systems for each.

We are very pleased to have Michael Becich and colleagues, Chenyu Li, Rumana Rashid, Eugene Sadhua, and Sandro Santagatac, contribute a new chapter on biorepositories. As Dr. Becich et al. maintain, the importance of biospecimens and their derivatives, particularly genomic sequencing and expression data coupled with deep clinical annotation from electronic health records, is fueling a new era of deep biologic interrogation of both the cell biology of human tissues and their diseased counterparts. The importance of computerized representation of both data and processes—including the formalization of roles and tasks—is underscored by Joyce Niland, Julie Hom, and Susan Hmwe in their chapter on Study Protocol Representation. The essence of any clinical study is the study protocol, an abstract concept that comprises a study's investigational plan and also a textual narrative documentation of a research study.

The section ends with an up-to-date look by Jeff Smith on the important policy issues concerning CRI. Smith provides deep detail on the history and implications of the *twenty-first Century Cures Act*, and he emphasizes that capitalizing on the numerous and extraordinary opportunities to improve development and delivery of new interventions will depend heavily on the application of CRI theory and methods.

Part II: Enabling Frameworks and Processes and Tools

Chapter 9: Data Sharing and Reuse of Health Data for Research

Chapter 10: Data Quality in Clinical Research

Chapter 11: Research Data Governance, Roles, and Infrastructure

Chapter 12: Informatics Approaches to Patient Recruitment

Chapter 13: Patient Registries for Clinical Research

Several chapters in this section cover the process of handling data in CRI and how to appropriately manage and use this data.

The use of clinical data for research is a tremendous challenge with perhaps the greatest potential for impact in all areas of clinical research. Standards specifications for the use of clinical data to populate research forms have evolved to support a number of very promising demonstrations of the "collect once, use many" paradigm. In her chapter on clinical data sharing and reuse of health data for research, Rebecca Kush covers various scenarios for data sharing, including who needs to share data and why. More importantly, she describes the history and future strategy of cooperation between major standards development organizations in health care and clinical research.

The quality of the data ultimately determines the usefulness of the study and applicability of the results. In "*Data Quality in Clinical Research*," Meredith Zozus, Michael Kahn, and Nicole Weiskopf address the idea that central to clinical research are data collection, quality, and management. They focus on various types of data collected (e.g., clinical observations, diagnoses) and the methods and tools for collecting these. Special attention is given to the development as use of case report forms (CRFs), historically the primary mechanism for data collection in clinical research but also the growing use of EHR data in clinical research and data bias. The chapter provides a theoretical framework for data quality in clinical research and also will serve as practical guidance. Moreover, Zozus et al. draw on the themes of workflows presented by Payne in the chapter on clinical research environment and advocate explicit processes dedicated to quality for all types of data collection and acquisition.

In the chapter on data governance, Anthony Solomonides provides many ways to look at how clinical research data is governed and the necessary steps to ensure data is managed and protected for research. In addition to providing key activities for proper data governance, he also explains various motivations and reasons supporting "why" effective data governance is important and ardently pursued. Dr. Solomonides describes organizational structures and processes that can be used to ensure data quality and patient and institutional protections. His chapter includes a case study.

Chunhua Weng and Peter Embi address information approaches to patient recruitment by dis-

cussing practical and theoretical issues related to patient recruitment for clinical trials, focusing on possible informatics applications to enhance recruitment. Their chapter highlights evolving methods for computer-based recruitment and eligibility determination, sociotechnical challenges in using new technologies and electronic data sources, and standardization efforts for knowledge representation.

Finally, and also related to patients, is a chapter on patient registries, provided by Rachel Richesson, Leon Rozenblit, Kendra Vehik, and Jimmy Tcheng. They highlight the growing importance of patient registries as curated data resources that can be enriched with multiple data types to support research, discovery, and healthcare quality improvement, thereby providing data infrastructure for learning health systems. Their discussion includes the scientific and technical issues for registries and reviews challenges and approaches for standardizing the data collected.

Part III: Managing Different Types of Data Across Clinical and Translational Research

Chapter 14: Best Practices for Research Data Management

Chapter 15: Patient-Reported Outcome Data

Chapter 16: Molecular, Genetic and other-omics Data

Chapter 17: Clinical Trials Registries, Results Databases, and Research Data Repositories

The premise of clinical research informatics is that the collection of data, and techniques for aggregating and sharing data with existing knowledge, can support discovery of new knowledge leading to scientific breakthroughs. The chapters that comprise this section are focused on state-of-the-art approaches to organizing or representing knowledge for retrieval purposes or use of advanced technologies to discover new knowledge and information where structured representation is not present or possible. A new chapter for this edition is on research data management by Anita Walden, Maryam Garza, and Luke Rasmussen. This chapter provides general guidance around the traditional and several forward-thinking aspects of clinical data management (CDM). It offers a foundational understanding of CDM and the scope of its activities. The chapter also describes the significance and relevance of CDM to researchers and data managers and summarizes best practices and resources for training and certification.

An important source of data, data reported by patients, is described thoroughly by Robert Morgan, Kavita Sail, and Laura Witte in the next chapter on "Patient-Reported Outcomes." The chapter describes the important role patient outcomes play in clinical research and the fundamentals of measurement theory and well-established techniques for valid and reliable collection of data regarding patient experiences.

This section also has a chapter on molecular, genetic, and other omic data that includes the increasing availability of genetic data that is becoming vital to clinical research and personalized medicine. The discussion provided by Stephane Meystre, Ramkiran Gouripeddi, and Alexander Alekseyenko primarily focuses on the relationship and interactions of voluminous molecular data with clinical research informatics, particularly in the context of the new (post) genomic era. The translational challenges in biological and genetic research, genotype–phenotype relations, and their impact on clinical trials are addressed in this chapter as well.

The full transparency of clinical research is a powerful strategy to diminish publication bias, increase accountability, avoid unnecessary duplication of research, advance research more efficiently, provide more reliable evidence (information) for diagnostic and therapeutic prescriptions, and regain public trust. Trial registration and results disclosure are considered powerful tools for achieving higher levels of transparency and accountability for clinical trials. New emphasis on knowledge sharing and growing demands for transparency in clinical research is contributing to a major paradigm shift in health research that is well underway. This section's final chapter by Karmela Krleža-Jerić, Mersiha Mahmić-Kaknjo, and Khaled El Emam discusses the use of trial registries and results databases in clinical research and decision-making. International standards of trial registration and their impact are discussed, as are the contribution of informatics experts to these efforts.

Part IV: Managing Different Types of Data Across Clinical and Translational Research

Chapter 18: Knowledge Representation and Ontologies

Chapter 19: Developing and Promoting Data Standards for Clinical Research

Chapter 20: Non-Hypothesis Driven Research: Data Mining and Knowledge Discovery

Chapter 21: Clinical Natural Language Processing in Secondary use of EHRs for Research

There is a natural appeal to ideas for transforming and exchanging heterogeneous data, which can be advanced using ontologies (or formal conceptual semantic representations of a domain). Kin Wah Fung and Olivier Bodenreider give us an overview of basic principles and challenges, all tied to examples of use of ontology in the clinical research space. This chapter covers the challenges related to knowledge representation in clinical research and how trends and issues in ontology design, use, and testing can support interoperability. Essential definitions are covered, as well as applications and other resources for development such as the semantic web. Additionally, major relevant efforts toward knowledge representation are reviewed. Specific ontologies relevant to clinical research are discussed, including the ontology for clinical trials and the ontology of biomedical investigation. Organizations, such as the National Center for Biomedical Ontology, that coordinate development, access, and organization of ontologies are discussed.

Rachel Richesson, Cecil Lynch, and W. Ed Hammond cover the topic of standards—a central topic and persistent challenge for informatics efforts. Their focus is on the standards development process and relevant standards developing organizations, including the Clinical Data Interchange Standards Consortium (CDISC) and Health Level Seven (HL7). They address the collaboration and harmonization between research data standards and clinical care data standards and discuss new standards initiatives, such as the HL7 Vulcan FHIR Accelerator initiative.

A chapter on nonhypothesis-driven research, authored by Mollie Cummins, Senthil Nachimuthu, Samir Abdelrahman, Julio Facelli, Ramkiran Gouripeddi, offers an overview of state-of-the-art data mining and knowledge discovery methods and tools as they apply to clinical research data. The vast amount of data that is warehoused across various clinical research enterprises, and the increasing desire to explore these to identify unforeseen patterns, require very advanced techniques. Examples of how nonhypothesis-driven research supported by advanced data mining, knowledge discovery algorithms, and statistical methods help elucidate the need for these tools to support clinical and translational research.

Much of the information in EHR systems is in the form of unstructured or free-text data, and as found in clinical notes and narratives, the ability to access these data has the potential or transform clinical research. As a final chapter in this section, authors Sunyang Fu, Andrew Wen, and Hongfang Liu discuss how new methods and tools are describe and how clinical natural language processing (NLP) has been adopted to computationally facilitate clinical research. Their chapter describes the foundation of clinical NLP and explains different NLP techniques that can be employed in the context of extracting and transforming narrative information in EHR to support clinical research.

Part V: Evolving Models and New Opportunities for the Transformation of Clinical Research

Chapter 22: Back to the Future: The Evolution of Pharmacovigilance in the Age of Digital Healthcare

Chapter 23: Evolving Opportunities and Challenges for Patients in Clinical Research

Chapter 24: Apps in Clinical Research

Chapter 25: Future Directions in Clinical Research Informatics

In this final section of the text, we also include topics that will continue to impact CRI into the future and that build upon the contexts, data sources, and information and knowledge management issues discussed in previous sections. Many of the topics included here are truly multidisciplinary and stand to potentially impact all clinical research studies.

Pharmacovigilance is an important and evolving discipline that is highly relevant to the future evolution of CRI application domains, particularly given the relevance of pharmacovigilance to patient safety and potential to impact population health. Informatics methods and applications are needed to ensure drug safety for patients and the ability to access, analyze, and interpret distributed clinical data across the globe to identify adverse drug events. Michael Ibara and Rachel Richesson provide a historical account of its evolution, as well as the increasing need for informatics methods and applications that can be employed to ensure greater patient safety. Various issues are explored in this context, including drug and device safety monitoring and changing paradigms and emerging infrastructures for detecting adverse drug events.

Two chapters in this final section tackle different perspectives on patients or consumers. Given the rapid advances in technology and parallel continued emphasis on patient empowerment and participation in decision making, Jim Andrews, Christina Eldredge, Janelle Applequist, and David Johnson consider the changing role of consumers in health care generally and in clinical research particularly. Traditional treatments of information behaviors and health communication are discussed, building to more current approaches and emerging models. Central to understanding the implications for clinical research are the evolving roles of consumers who are more engaged in their own decision-making and care and who help drive research agendas. The tools and processes that support patient decision-making, engagement, and leadership in research are also briefly described here.

A new chapter, "Apps in Clinical Research," is also included in this edition. In this chapter, Brian Douthit describes how the introduction of "apps" (s*oftware applications that can be installed and run on computers, tablets, or smartphones*) are changing the landscape of clinical research by opening new avenues for administrating and evaluating interventions. The almost ubiquitous use of smartphones and apps has transformed many industries, and a similar transformation of health care and clinical research appears inevitable. In this chapter, Dr. Douthit outlines the major events leading to today's app infrastructure, current uses of apps in clinical research, design considerations, and future directions. The book concludes with a brief chapter by Peter Embi summarizing the challenges CRI researchers and practitioners will continue to face as the field evolves and new challenges arise. This concluding chapter helps in envisioning the future of the domain of clinical research informatics. In addition to outlining likely new settings and trends in research conduct and funding, the author cogitates on the future of the informatics infrastructure and the professional workforce training and education needs. A focus of this chapter is the description of how clinical research (and supporting informatics) fits into a bigger vision of a learning health system and of the relationship between clinical research, evidence-based medicine, evidence-generating medicine, and quality of care.

Conclusion

The overall goal of this book is to contribute to the ongoing discourse among researchers and practitioners in CRI as they continue to rise to the challenges of a dynamic and evolving clinical research environment. CRI is an exciting and broad domain, leaving ample room for future additions or other texts exploring these topics more deeply or comprehensively. Most certainly, the development of CRI as a subdiscipline of informatics and a professional practice area will drive a growing pool of scientific literature based on original CRI research, and high-impact tools and systems will be developed. It is also certain that CRI groups will continue to support and create communities of discourse that will address much needed practice standards in CRI, data standards in clinical research, policy issues, educational standards, and instructional resources.

The scholars that have contributed to this book are among the most active and engaged in the CRI domain, and we feel they have provided an excellent starting point for deeper explorations into this emerging discipline. While we

have by no means exhausted the range of topics in this new edition of CRI, we hope that readers will see several themes stand out throughout this text. The authors illustrate well that the CRI domain can keep up with the pace of innovation even in hard times (like pandemics). It is likely that moving forward, the use of informatics and computing will continue to evolve to address emergent research needs and technical capabilitie and accelerate the generation of new knowledge and insight to guide the course of human and global evolution in ways we cannot even predict.

References

1. Mayer D. A brief history of medicine and statistics. In: Essential evidence-based medicine. Cambridge: Cambridge University Press; 2004. p. 1–8.
2. Atkins HJ. The three pillars of clinical research. Br Med J. 1958;2(5112):1547–53.
3. Bacchieri A, Della CG. Fundamentals of clinical research: bridging medicine, statistics and operations, statistics for biology and health. Milan: Springer; 2007.
4. Friedman CP. A "fundamental theorem" of biomedical informatics. J Am Med Inform Assoc. 2009;16(2):169–70.
5. NIH. The NIH Director's panel on clinical research report to the advisory committee to the NIH director. 1997. http://www.oenb.at/de/img/executive_summary%2D%2Dnih_directors_panel_on_clinical_research_report_12_97_tcm14-48582.pdf. Accessed 2011.
6. Zhu N, Zhang D, Wang W, et al. A novel coronavirus from patients with pneumonia in China, 2019. N Engl J Med. 2020;382(8):727–33.
7. Cohen J. With record-setting speed, vaccinemakers take their first shots at the new coronavirus. Science. 2020. https://www.science.org/content/article/record-setting-speed-vaccine-makers-take-their-first-shots-new-coronavirus. Accessed 09 Oct 2022.
8. FDA. https://www.fda.gov/science-research/science-and-research-special-topics/real-world-evidence.
9. FDA. Conduct of clinical trials of medical products during the COVID-19 public health emergency. https://www.fda.gov/media/136238/download.
10. O'Brien EC, Sugarman J, Weinfurt KP, Larson EB, Heagerty PJ, Hernandez AF, Curtis LH. The impact of COVID-19 on pragmatic clinical trials: lessons learned from the NIH Health Care Systems Research Collaboratory. Trials. 2022;23(1):424. https://doi.org/10.1186/s13063-022-06385-8.
11. Valenta AL, Meagher EA, Tachinardi U, Starren J. Core informatics competencies for clinical and translational scientists: what do our customers and collaborators need to know? J Am Med Inform Assoc. 2016;23(4):835–9. https://doi.org/10.1093/jamia/ocw047.
12. Embi PJ, Payne PR. Clinical research informatics: challenges, opportunities and definition for an emerging domain. J Am Med Inform Assoc. 2009;16(3):316–27.
13. Sanchez-Pinto L, Mosa ASM, Fultz-Hollis K, Tachinardi U, Barnett WK, Embi PJ. The emerging role of the chief research informatics officer in academic health centers. Appl Clin Inform. 2017;8(3):845–53.
14. Solomonides A. Review of clinical research informatics. Yearb Med Inform. 2020;29(1):193–202.
15. Forrest WH, Bellville JW. The use of computers in clinical trials. Br J Anaesth. 1967;39:311.
16. Collen MF. Clinical research databases—a historical review. J Med Syst. 1990;14(6):323–44.
17. Pryor DB, et al. Features of TMR for a successful clinical and research database. In: Blum BI, editor. Proceedings of the sixth annual symposium on computer applications in medical care. New York: IEEE; 1982.
18. Blois MS. Information and medicine: the nature of medical descriptions. Berkeley: University of California Press; 1984.
19. Blois MS. Medical records and clinical databases: what is the difference? MD Comput Comput Med Pract. 1984;1(3):24–8.
20. Breitfeld PP, et al. Pilot study of a point-of-use decision support tool for cancer clinical trials eligibility. J Am Med Inform Assoc. 1999;6(6):466–77.
21. Carlson R, et al. Computer-based screening of patients with HIV/AIDS for clinical-trial eligibility. Online J Curr Clin Trials. 1995;4:179.
22. Mansour EG. Barriers to clinical trials. Part III: knowledge and attitudes of health care providers. Cancer. 1994;74(9 Suppl):2672–5.
23. ICH. Guideline for good clinical practice E6(R1). International Conference on Harmonisation. 1996.
24. Shortliffe EH, Cimino JJ. Biomedical informatics: computer applications in health care and biomedicine health informatics. New York: Springer Science+Business Media, LLC; 2021.
25. Hersh WR. Health informatics: practical guide. 8th ed. Morrisville: Lulu.com; 2022. p. 460.
26. Embi P. AMIA CRI years in review. 2018. http://www.embi.net/cri-years-in-review.html. Accessed 1 July 2018.
27. Bacchieri A, Cioppa G. Fundamentals of clinical research: bridging medicine, statistics and operations. Milan: Springer; 2007.