I. SPECIFIC AIMS

New biomedical discoveries, advances in information technology, and implementation research offer the potential for immediate, continuous, and transformative improvement in health care, but not within our current healthcare system. As envisioned by the Institute of Medicine,[1] a learning healthcare system comprises a community of front-line clinicians, patients, and scientists who view each clinical encounter as an opportunity to learn and to improve patient outcomes. They rely on large registries and open-science networks that foster collaborative improvement, research, data sharing, and innovation.[2-4] The registry captures, manages, and provides access to condition-specific information in a uniform way for a list of specific patients to support their clinical care.[5] In its most advanced state, a learning healthcare system combines comparative effectiveness (CE) research (T2 translation) with quality improvement (QI) science (T3 translation) to advance patient outcomes.[6] **We propose to achieve this vision of a learning healthcare system by using open-source informatics tools to enhance an existing registry for children with inflammatory bowel disease (IBD) to enable better standardization and customization of care and CE research to create new knowledge about what care is best for which patients.**

As part of federal initiatives to speed the introduction and "meaningful use" of health information technology nationally, the Centers for Medicare & Medicaid Services provides incentive payments to providers that electronically capture clinical data from patients to improve quality, outcomes, and efficiency.[7] Electronic health record (EHR) systems have the potential to become the tool for learning at the point-of-care, although most are patient-focused and do not natively support the population management functionality required by a patient registry. The national transition to EHRs lends urgency to the tasks of developing software and processes to populate registries with data entered once (i.e., single data entry for clinical care, QI, and CE research). Recent advances in data sharing technologies using open-source tools[8-11] and new methods for network-based research and innovation provide opportunities for creating a national, distributed, pediatric data network based upon capture of EHR data into registries.

We propose to build upon existing open-source software,[9, 11, 12] with which we have extensive experience, to create an enhanced registry that is modular, versatile (i.e., can support longitudinal comparative effectiveness research and advanced QI) and scalable to many chronic conditions that affect children (a priority population). The Pediatric EHR Data Sharing Network (PEDSNet), the governance institution, will manage the data network and be developed to ensure its sustainability and growth as an open-science community. We will test our ability to provide optimal support of QI and CE research (an implementation sciences agenda[13]) within the successful ImproveCareNow practice-based research and improvement initiative.[14, 15] Currently, all ImproveCareNow network sites collect clinical care data using structured, paper-based clinical encounter forms. Data are then entered into a centralized clinical registry via a web-based interface (i.e., sequential entry of clinical data into 2 sources). We will change this laborious, error-prone process to collect data once at the point of care. Our aims are to:

Aim 1: Enhance an existing registry to support a learning healthcare system for pediatric inflammatory bowel disease by capturing needed data directly from electronic health records, improving the quality of collected data using new tools we have developed for recording clinical information during a patient encounter, and facilitating interventions to improve the quality of care for children.
Aim 2: Use Quality Improvement methods to implement enhanced IBD-registry features to enable management of IBD care center patient populations and to increase patient participation in care.
Aim 3: Use data from the enhanced registry to compare the effectiveness of alternative treatment strategies for pediatric Crohn's disease (CD) patients, with a special focus on timing of biologic agents.
Aim 4: Develop governance structures for the network that engages patients and provides oversight of privacy, confidentiality, and data access, as well as scientific and technical concerns.

The project will form a unique community of children and youth with IBD, their families, pediatric IBD clinicians, informaticians, QI specialists, and CE researchers who will work together to improve patient outcomes using learning healthcare system principles. Our long-term goal is to develop an approach and set of tools to extend this work to other chronic disease communities devoted to advancing the health of children.

II. RESEARCH STRATEGY

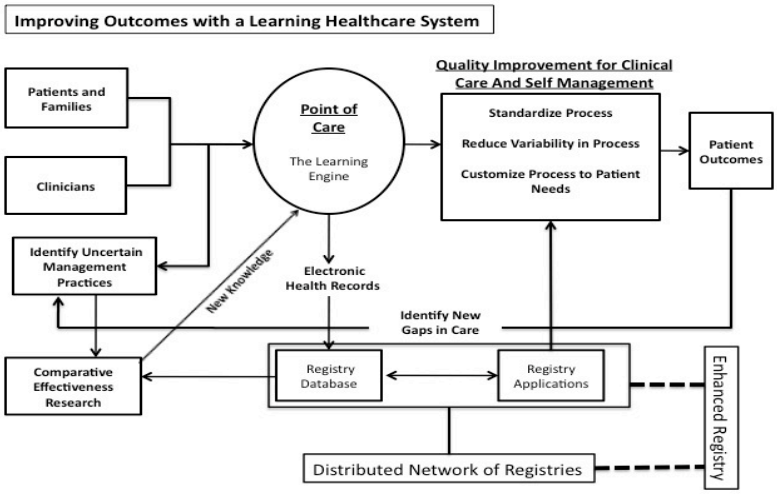
SIGNIFICANCE

A BACKGROUND

A1 Learning Healthcare Systems: A Model for Improving Patient Outcomes with QI and CE Research

Through its EHR Innovation Collaboratives, the Institute of Medicine (IOM) has stimulated the development of several learning healthcare systems; the Pediatric EHR Data Sharing Network (PEDSNet) is one of these organizations (see McGinnis Letter of Support). Core concepts of the IOM vision include: a focus on continuously improving outcomes; learning as a partnership enterprise among patients, clinicians, and researchers; the point-of-care serves as the knowledge engine; advancing clinical data as a public utility; building CE research into practice; and, a governance model that promotes diverse leadership.[16] Although these concepts are laudable and foundational to the formation of a learning healthcare system, there is no blueprint that describes the approach, processes, or organizational architecture necessary for achieving the vision.

For the proposed project, we intend to operationalize the IOM's vision for the learning healthcare system. This is depicted in the Figure. Patient/families and clinicians, with the assistance of CE and QI researchers, identify management practices for which there is sufficient evidence to guide specific decision-making and others that merit scientific investigation. Clinical and self-management practices are standardized, made reliable, and customized to patients' needs using QI methods. Standardization and high reliability enhances signal detection in research studies, which is an often-overlooked link between CE research and QI. Both CE research and QI are made possible using digitized data from electronic health records, which is stored in the registry database. Informatics applications are developed to manipulate these data to facilitate the QI. Outcomes are continuously monitored to identify new gaps in care. These processes form an iterative cycle--data collected at the point of care using EHRs, new knowledge generated and evaluated (i.e., learning using CE research methods), and management practices standardized with variation across providers minimized. A key innovation of the approach we propose is the linkage of disease-specific learning systems into a broader pediatrics open-science learning community via PEDSNet, concurrently creating a distributed network of registries (each disease registry can itself be a network of local registries, so this distributed network is, in essence, a network of networks). The combination of EHR-based registry data suitable for learning, informatics applications to enable QI, and linkage to a network of registries constitutes the project's "enhanced registry."



Several scientific advances are converging to facilitate the development of our proposal for an open-science learning healthcare system. Research has become more networked and data-intensive, trends fueled by inexpensive data storage, high computation processing, and a global high speed internet-based communication infrastructure. Although researchers freely share knowledge through peer-reviewed publications, data sharing has been much less common. In a clinical data-sharing environment, governance is needed to address human subjects concerns, conflicts of interest and dispute resolution, and to overcome legal, ethical, and social obstacles.[17] In this proposal, PEDSNet will serve this role. Although there are few examples of open-access data sharing in healthcare, open access to genomics data provides a model worthy of consideration.[2, 3] Early release of the draft human genome sequence led to 30 disease-gene associations even before the manuscript describing the full genomic dataset was in print.[18] Most researchers embrace data sharing conceptually; yet, actual examples are few. Scientists are concerned about being scooped or data being misused,[19] and institutions worry about sharing information that can harm competitiveness.[20]

A legal system that was crafted in an "analog era" poses challenges that hinder the reuse of knowledge

and slows innovation in the digital age.[21, 22] Much has been written about the negative impact of HIPAA legislation, such as its lack of true privacy protection when data are digitized,[23] and there are substantial concerns about the additional privacy and notification restrictions in the ARRA HITECH legislation.[24] Novel approaches to human subjects protection, intellectual property, and the creation of appropriate academic incentives for collaboration are evolving from organizations such as Creative Commons, [25] which provides creators various copyright license options to make their work available to the public on generous terms. Similar principles are being applied to scientific sharing through the work of Science Commons.[26]

In collaboration with Science Commons, we will assess and develop solutions for the legal, ethical and social barriers to promoting and facilitating collaboration and sharing of data and related intellectual resources in a pediatric, open-science learning network (see **Aim 4**). Clearly articulating these challenges and developing solutions to them are essential to the success of PEDSNet, which aims to collect data from many institutions nationally and create a federated "data trust" to facilitate creation of registries and research on children with chronic disease. We expect that these solutions will be important to many other similar networks.

A2 Electronic Health Records, Registries, CE and QI Research

Electronic patient registries are the substrate for outcomes improvement in a learning healthcare system. They are one of the better established types of clinical information systems, having been first implemented in the late 1980's.[27] Registries are population-based healthcare tools that are useful for management of chronic disease, product and drug safety monitoring, natural history of disease studies, and improving quality.[28] Although they have been in existence for several decades, many suffer from a number of technical limitations that prevent them from being widely adopted or easily integrated into clinical workflows. For example, most registries are developed to address a specific clinical problem and therefore are essentially "one-of-a-kind" creations custom-designed for a specific project. The development of a second registry, even in a closely related clinical field, requires roughly the same amount of effort as creation of the original registry. Second, it is challenging to feed data directly from an EHR into a registry. A minority of registries can upload data that have been extracted from an EHR into a flat file. Third, many registries have limited reporting capabilities. Feedback is generally provided in the form of monthly quality and population reports and there is limited ability to explore underlying data. Fourth, most are "centralized," meaning all data are sent to one site. If a participating site withdraws from the registry, it can be difficult to obtain its own data and ensure that it is no longer included in the overall registry. If the central site withdraws, the entire registry can collapse. Fifth, most EHRs are patient-focused by design and do not natively support the type of functionality required by a population focused registry. Performing population-level queries, for instance, is difficult, if not impossible. Sixth, vendor-developed EHRs are proprietary and usually not open to modification by their customers. This makes it difficult to add new data sources and data types, integrate data from outside systems or institutions, or associate the registry data with bio-repositories. Lastly, unless institutions are within the same health system, there is little interoperability among EHRs. These and other challenges to disease registries were summarized in a recent NIH conference on patient registries and biospecimen repositories, for which co-I Forrest served on the planning committee.[29]

To overcome these technical obstacles to advancing registries, we propose (**Aim 1**) to develop an infrastructure that supports EHR-linked registries, designed according to the following principles:

1. Data entry efficiency: Data are entered once at the point of care and are transmitted into the registry. This removes the typical registry requirement of having to perform double data entry or abstract all of the information from a clinical note. Processes are put in place to monitor data quality.
2. Flexible EHR linkage: The enhanced registry is designed to be linked to the EHR, although an EHR is not required. Data can flow into the system through multiple inputs, allowing institutions with varying degrees of IT expertise and EHR maturity to participate.
3. Distributed network: Hospitals and other health care providers are concerned about protecting patient data and are becoming very cautious when it comes to allowing patient data to be managed by an outside institution. Distributed networks allow data holders to retain their data, thereby providing local control over the exchange of clinical information. This strategy helps mitigate several privacy and legal concerns when human subjects information is shared between institutions.
4. Modular application layer: To facilitate QI and population management, an open-source application layer is built using standard web services. This will support conventional functions provided by registries such as visit planning tools, population and QI reports and across site comparisons, and is

modular in nature to allow new capabilities to be added without re-engineering the system. This application-agnostic, platform-oriented approach is similar to the iPhone. The registry provides the basic framework, while the open-source software community is free to develop new apps or plug-ins.

5. Governance: The registry is supported by a strong system of governance, to oversee, sustain and scale this approach.

Toward Distributed Data Networks: Transition to electronic health records nationally provides new opportunities for populating registries with data obtain from routine clinical workflows. Recent advances in data sharing methodologies using open-source tools,[8-11] and novel methods for network-based research and innovation facilitate the creation of a national distributed pediatric data network based upon capture of EHR data into registries, which themselves are linked to federated data warehouses. Creating a master database that is retained at a data-coordinating center is the conventional approach to sharing research data across institutions. This model allows for centralized data governance, data quality monitoring, and analytic services. While effective for single studies, the centralized data model is expensive to maintain, does not promote secondary uses by the broader scientific community, inhibits an open-science approach to learning, and does not leverage the growing presence of EHRs.[30] Due to significant advances in Internet and data sharing technologies, distributed research networks are gradually replacing the central data center model. A small, but growing, number of distributed research networks have been implemented across a spectrum of clinical conditions and practice settings.[31-34] Within EHRs, however, syntactic interoperability (the ability to exchange data) remains a challenge. Open-source solutions for health information exchange such as CONNECT, which uses Nationwide Health Information Network (NHIN) standards and governance, are in only the nascent stage.[35] Even if (and when) the technical barriers to EHR-based communication are resolved, enabling semantic interoperability between different EHR vendors remains an outstanding issue.[28] There is little consensus on how to describe and label EHR data. Efforts by CDISC (http://www.cdisc.org), with projects like BRIDG and SHARE, aim to solve some of these challenges, but it remains to be seen whether they will achieve meaningful adoption in clinical operations. Relying on a distributed network removes the need for semantic interoperability between EHRs, because that translation happens outside the EHR.

Strategies for Improving Data Quality: Simultaneous advances in EHRs' functionality, graphical user interface designs, and mobile/wireless hardware have expanded the sophistication of tools that support direct clinician entry of detailed clinical data as discrete coded data elements.[36, 37] Although currently optimized for clinician productivity over data accuracy, EHR interfaces have many of the real-time data validation capabilities that are found in dedicated clinical trials electronic data entry systems—e.g., range checks and mandatory entry requirements. In addition to capturing more clinical data in coded format directly at the point of care, natural language processing can extract information from dictated reports and notes, such as discharge and procedure notes, pathology and radiology reports, and even death certificates.[38] The vision of capturing high-quality clinical data during standard clinical care (**data-in once**) that can be shared over a secure, HIPAA-compliant collaborative research network for a wide range of secondary uses, including CE research and continuous clinical QI, is clearly at hand. Nonetheless, evidence exists that current EHR documentation practices may not result in complete, consistent data collection.[39-43] Given the diversity of EHR systems and different uses of clinical documentation tools, this project will address the strengths and weaknesses of a variety of data capture methods to achieve "research grade" data obtained during routine clinical care.

A3 Sub-Specialty Research and Improvement Networks: "Laboratories" for QI and CE Research

Most pediatric chronic conditions meet the NIH definition for a rare disease,[44] and no single care center has a sufficient number of patients to produce generalizable knowledge, a barrier that, unless addressed by networks, will slow the pace of knowledge acquisition and outcomes improvement. The American Board of Pediatrics (ABP) has supported the establishment of sub-specialty improvement and research networks in all 13 pediatric sub-specialties (see Letter of Support from ABP, Paul V. Miles). This effort extends the successes of pediatric clinical networks that use data for research and (increasingly) improvement. Examples include: (1) Vermont Oxford Network (VON) (est. 1988) that is dedicated to improving the quality of neonatal ICU care[45] and (2) the Children's Oncology Group (COG) (est. 1998) that focuses on clinical trials of new therapies, as well as studies of how to improve the delivery of existing therapies for pediatric cancer. The dramatic improvement in cancer survival rates from <10% in 1962 to over 80% today is a reflection of the benefits of an infrastructure that can systematically standardize care and enroll patients in clinical trails.[46] The common themes drawn from collaborative pediatric research networks that have demonstrated marked improvement in

the outcomes of children with chronic disease are an unrelenting commitment to collecting high quality data, continuously evaluating and proving their value to clinicians making in-the-trenches decisions, and the long-term engagement of the participants and their institutions to sustaining the network.

Sub-specialty research and improvement networks offer advantages that are foundational for research "laboratories"[47] that can support a pediatric chronic disease learning system. Creation of total population registries at each site provides large and diverse study samples. By standardizing practice, they reduce variations in outcomes due to care delivery, thereby increasing statistical power. By linking research to care delivery and engaging clinicians directly, they provide a forum for user-led CE research,[48] a core attribute of the learning healthcare systems.[16] Not only are the end-users of comparative effectiveness research--i.e., clinicians--in a unique position to identify critical healthcare knowledge gaps, they along with their patients are the final common pathways for change at the point-of-care. Since 2004, 8 of the 13 pediatric sub-specialties have begun the development of such improvement and research networks. Pediatric gastroenterologists formed ImproveCareNow (which is part of the proposed project) to study and improve the care of inflammatory bowel disease. Despite the pressing need to scale up these networks, expansion is hampered by the lack of an informatics infrastructure capable of supporting needed expansions while reducing the costs of conducting research. Sub-specialty research networks typically collect data manually, using specially designed disease-specific forms. As a result, they do not capture data directly in electronic health records or incorporate the data into federated data warehouses. Furthermore, they do not make the data available to other researchers, allowing it to be used to address a wide variety of secondary questions. These methods are costly, ad hoc (i.e., depend on grant funding), and slow the pace of discovery by siloing datasets. We will address these needs by developing an enhanced registry **(Aim 1**) that permits capture of research-grade data directly from EHRs and incorporates it into a federated research data warehouse. The infrastructure is versatile, built with open-source tools, and generalizable, which will make it attractive to many sub-specialties. This project will be a prototype for other subspecialty research networks that can benefit from this effort with substantial cost savings. In addition to these technical advances, we will work with Creative Commons and our collaborators within PEDSNet to increase their willingness to share data and best practices to create learning communities.

A4 Defining Optimal Treatment Strategies for Pediatric Crohn's Disease (CD)

CD is a type of inflammatory bowel disease (IBD) that is associated with significant medical, psychosocial, and economic impact. Childhood onset disease occurs in approximately 25% of cases and is particularly aggressive. Children experience significant psychosocial impact.[49] Surgical intervention is required in 17% of cases within five years and 28% within 10 years of diagnosis.[50] Over the last several decades, therapeutic advances in the treatment of pediatric CD have included the widespread use of immunomodulators such as 6 mercaptopurine (6-MP),[51] azathioprine, and methotrexate.[52] More recently, the anti-tumor necrosis factor (anti-TNFα) biological agents have been added to the therapeutic armamentarium.[53] Despite these advances, there has been little or no improvement in outcomes over the past four decades.[54] Studies have yet to define strategies for how, when, and in whom to initiate these treatments. Lack of evidence on how to best treat the condition is highlighted by widespread practice variation.[55] Variation in practice and the risks and costs of various treatment strategies are an important reason why timing of biologic agents was selected as **one of the IOM's top 25 CE research priorities**. This IOM CE research objective on timing of biologics suggests the need for research on sequential patterns of treatment decisions with each decision dependent on the prior decision and treatment response (i.e., a treatment strategy). Current evidence is based on comparisons of one treatment with another, rather than assessing the optimal dynamic treatment strategy.[56]

No studies directly comparing these treatment options have been performed in children with CD. It is widely believed that anti-TNFα therapy is more efficacious than thiopurines, owing largely to the fact that these agents are effective in inducing and maintaining remissions in patients who failed to respond to thiopurine therapy. Their use remains limited, however, by concerns about cost, safety and the lack of data about which patients will benefit most from such therapy. The cost of one year of anti-TNFα therapy can be upwards of $50,000, as compared with approximately $3,000 for corticosteroids and thiopurines. Anti-TNFα's have been considered high-risk medications, with the major potential toxicity being serious infections and malignancy, although the evidence for these concerns has been called into question.[57-59]

One reason for proposing this CE research objective is that the best approach to timing of biologics is unknown. There are two dominant theories in pediatrics. The *step-up* approach involves initial treatment of patients with steroids and thiopurines at diagnosis or very early in the course of disease.[60] If the response is inadequate, anti-TNFα therapy is begun early (<6 months) or later (>6 months). The rationale for early step-up

treatment is that pediatric CD is believed to be aggressive (more so than the adult-onset variant).[61, 62] On the other hand, a randomized controlled trial by Markowitz and colleagues suggested that the combination of corticosteroids and thiopurines initiated at diagnosis could achieve sustained, steroid free remission in 91% of pediatric CD patients,[51] as compared to remission rates of approximately 50% in those treated without thiopurines. Clinicians commonly cite these data as the justification for the later step-up approach. The Markowitz study, however, has been called into question because just 55 patients comprised the sample (the proposed study in this application will collect data on 2,700 patients). The "rapid step-up" approach has been modified because of advances in pharmacogenetics. The ability to measure thiopurine methyltransferase activity and monitoring of thiopurine metabolites has enabled gastroenterologists to start patients on higher doses of thiopurines safely, enabling them to increase the dose of medication rapidly to optimal levels. It remains unclear whether more rapid achievement of an appropriate dose results in improved patient outcomes.

An alternative to the step-up therapy is what has been termed the "top-down" approach: initial therapy with anti-TNFα biologic agents prior to corticosteroids and thiopurines. Proponents of this approach believe that early therapy with biologics will modify the natural history of this chronic disease. This treatment strategy is based on limited evidence from two small studies conducted in adults.[63, 64] Earlier use of anti-TNFs is also supported by sub-analysis of clinical trials, which showed an improvement in important secondary outcomes, including improved growth,[53] fewer hospitalizations,[65] and less need for surgery.[65]

Although pediatric CD is the most significant and debilitating gastrointestinal illness in childhood, there is extremely limited information to guide choice of a treatment strategy that is optimal for a given patient. Research that estimates accurately the clinical effectiveness of existing treatment strategies in real-world settings and compares the effectiveness of such approaches in large numbers of children is critically needed to inform clinical practice. The proposed study (Aim 3) will fill these knowledge gaps by conducting the largest outcomes study to date on pediatric CD patients (n=2,700).

INNOVATION AND APPROACH

B PRELIMINARY STUDIES

Our team joins investigators from multiple institutions with complementary expertise in informatics, practice-based research and improvement, effectiveness research, patient-reported outcomes, and causal modeling using observational data. The following preliminary studies indicate we have succeeded in overcoming the research infrastructure, informatics, and human systems barriers that have slowed progress towards CE and quality improvement research. Building on our past successes, we are well positioned to advance the state of the field with the proposed project.

B1 Existing Registries and Open-Source Informatics Tools

Currently, all ImproveCareNow network sites collect data using structured clinical encounter forms. Data are then entered into a centralized clinical registry via a web-based interface (i.e., sequential entry of clinical data into 2 sources). After a site joins ImproveCareNow, it adapts its encounter forms (using examples from other sites) to structure them to capture all required data elements, each of which has a standardized operational definition (e.g., current medications are defined as existing and new medicines a patient is prescribed at the end of the encounter [rather than the beginning of the encounter]), and to make them feasible within their setting. (Examples of process flow diagrams and data capture auditing reports are included in the Appendix 2, EHR-based Capture of Registry Outcome Measures). The adaptation and testing process is supported by ImproveCareNow quality improvement training and coaching as part of the implementation process. Data entry at present is a laborious process and does not utilize the advantages that linkage of the disease specific registry with an electronic health record could provide. High quality data are fundamental to the ImproveCareNow's ability to improve outcomes of care, but present methods are neither scalable nor sustainable. More accurate data make it easier for practices to see the effects of their improvement efforts over time, facilitates sites' ability to manage patients as a population, and enhances the network's ability to learn from differences in performance across sites. The uses of data for these purposes represent key elements of the concept of "meaningful use"[22] of health information technology, specifically electronic health records.

A key design goal of this proposal is to apply the principles of quality and systems improvement to move ImproveCareNow members to enhanced registries that collect data from local EHRs utilized at the point of care during routine clinical workflows and can then be federated with other registries within a collaborative

learning research network. The enhanced registry will use public domain informatics tools that can be distributed and implemented widely at substantially reduced costs to participating sites and reused for multiple diseases and specialties. From an informatics perspective, the enhanced registry architecture includes the following components: a) linkage to the EHR; b) an application layer that allows the development of functionality to support QI and CE research; and, c) the ability to perform federated queries among a network of local registries, allowing for the creation of a larger "virtual" registry.

This vision of an enhanced registry is instantiated by combining elements of the i2b2 informatics framework,[8, 9] the SHRINE federated query architecture,[11] and application modules, in the form of an i2b2 toolkit, developed by our team (described below). The i2b2 data warehouse is used to combine clinical information linked from the EHR into an integrated data repository (component a). Our i2b2 toolkit provides the necessary functionality in a modular fashion, allowing users to choose the capabilities that are appropriate for their project (component b). The SHRINE federated query architecture can be used to link local registries into a larger network, allowing for queries on a larger, more diverse patient set (component c). Such a design gives enhanced registries the ability to capture data directly from the EHR and other sources, integrate it into a single repository, and expose it for both population and patient management, as well as other analyses. The modular construction allows the enhanced registries to support the diverse needs of QI and CE research.

The i2b2 informatics framework will serve as the foundation for our registry. Funded by the NIH's National Center for Biomedical Computing at Harvard Partners, i2b2 was developed to meet the needs of clinical and translational researchers desiring an interoperable, open-source, scalable and publicly available data framework for clinical research.[9] An Oracle or Microsoft SQL database can serve as a repository of data for i2b2.[12] The architecture offers tremendous flexibility in mapping data in varying formats from disparate sources into a unified representation, allowing investigators to easily add new data elements using conventional fact-centered star schema, customizing data flows, and creating applications for aggregating and analyzing the data.[8] At this time, 30 Academic Health Centers have chosen i2b2 for their internal research purposes, including half of the currently awarded CTSAs and six international sites. Several industries, including major pharmaceutical companies (Johnson and Johnson) have also adopted this platform for various specialized applications (See Letter of Support from Isaac Kohane, Director, i2b2 National Center for Biomedical Computing). The philosophical underpinning of i2b2 is that data sharing is at the core of all science, which is consonant with the principles of the open-science PEDSNet research community.

Importantly, for clinical research, data stored in i2b2 are organized in a way to protect data and patient privacy, even when these data are shared across organizations. The Children's Hospitals of Cincinnati and Denver have large production-scale implementations of i2b2 supported by the institution, which offer the necessary implementation, deployment, and maintenance knowledge of the platform for this project. Furthermore, the i2b2 architecture has been extended seamlessly to link data obtained in EHRs to i2b2 research warehouses and several new software applications relevant to the proposed project have been developed. For example, Dr. Marsolo (co-I) has been funded as part of two NIH-funded initiatives to improve the ability of i2b2 to support general-purpose clinical research: (1) a multi-institutional registry to study pediatric rheumatic diseases (PI: Laura E. Schanberg, RC2-AR058934-01, CARRA: Accelerating Toward an Evidence Based Culture in Pediatric Rheumatology) and (2) a national registry for patients with eosinophilic esophagitis (PI: Marc E. Rothenberg, R01-DK076893, Comparative Effectiveness of Pediatric Eosinophilic Esophagitis Registry). As part of these NIH-funded efforts, the following software modules are available and will play a critical role in ensuring the development of the modular enhanced registry (also see the Appendix 3, Technical).

1. Web-based forms for data capture and display – Enable users to configure web forms using a template based on the CDISC ODM standard.[66] These forms can then display data from i2b2 in a manner familiar to those who work with electronic data capture systems or other clinical research tools. These forms also allow direct data entry into i2b2, remove the need for users to re-enter clinical data.
2. Reporting and visualization tools – Give users advanced charting capabilities within the i2b2 workbench and mechanisms for visualizing patient outcomes. These capabilities also enable population level queries—i.e., a single site can chart its progress against the aggregate group.
3. File upload and validation – Permits user to upload data files and match them to an i2b2 ontology. This tool provides detailed error checking at the time of upload using range checks, logic rules, and detailed audit trails. It has been designed specifically to enable high quality data management consistent with Society of Clinical Data Management Good Clinical Data Management Processes.[67]

4. Data exporting tools –The very limited i2b2 data exporting capabilities are being extended with a new export function that allows users to extract data in a "spreadsheet-like" format.

The code for these applications is currently being made publicly available through the i2b2 academic users group (See Letter of Support from Isaac Kohane, Director, i2b2 National Center for Biomedical Computing), highlighting our commitment to contributing innovative tools to the research community.

B2 Design of Large-Scale Research and Improvement Collaborative Pediatric Networks

Starting in 2003, Dr. Margolis led the design of the Board of Pediatrics sub-specialty network initiative, spawning networks in 8 of the 13 pediatric sub-specialties.[4] Using a structured design process based on principles of complex system development,[68] the approach included explicit delineation of users' needs (i.e., specialty physicians and care teams), which resulted in a design with three core features: (1) collaborative networks that could achieve large Ns to ensure adequate generalizablilty; (2) methods to enable collection of clinical data for research and improvement as part of routine clinical care; and, (3) support to standardize care using QI methods that reduce unwanted provider-specific variations in care. Dr. Margolis continues to extend this theme of research as co-PI of an NIH "Transformative RO1" (R01DK085719-01) aimed at the development of collaborative chronic illness care networks that engage physicians, patients and researchers simultaneously.

Particularly relevant for this proposal, the network design recognized the need to collect data (**once**) during routine clinical care that could also be used to evaluate the comparative effects of process variations on outcomes of care. The design embedded methods that start with the desired end-result, then works backward to front-line data collection and data flow, an approach advocated by James[69] and others.[45] This structured approach collects a parsimonious set of data that include only those elements that lead to actionable reports while comprehensively covering an entire process of care. The approach has enabled care teams to integrate high quality data collection into workflow in each of the collaborative networks and at Cincinnati Children's Hospital Medical Center (CCHMC) for >60 conditions (See Appendix 2 for a description of this process).

The design process identified six attributes necessary to enable learning from every patient: (1) a structured process to developing measures; (2) a variety of ways to move data in/out of the system, ensuring adaptability to variations in work processes; (3) standards for content and messaging; (4) standardized and user-configurable reports; (5) assurance of data quality, including data validity checks and skip checks, auditing features to track data entry, change and deletion events, measures to assess data completeness and procedures for error correction; and, (6) security and confidentiality, including features to guard data integrity, and policies and procedures to control and audit access to protected health information. These design principles are guiding the ongoing development of the enhanced registry.

B3 ImproveCareNow: A Successful Practice-Based Research and Improvement Network

CoI-s Colletti, Crandall, Kappelman, and Margolis lead the ImproveCareNow network (formerly the PIBDNet Trailblazer Improvement Collaborative), which was created under the sponsorship of the American Board of Pediatrics (ABP). Its mission is to improve the outcomes of care for children with IBD. The network was established in 2007, guided by the vision of the ABP, to facilitate practice-based improvement activities that will enable pediatric gastroenterologists to meet new competency requirements in systems thinking and performance in practice. It has since grown to 23 sites located at universities, children's hospitals, multi-specialty clinics, and private practices, and includes both rural and urban settings nationally.

Initial activities of the network concentrated on measuring, standardizing and improving care delivery and outcomes using a complete population registry of patients seen at each site. Network improvement activities are managed using a modified Breakthrough Series[TM-70] method and Wagner's improving chronic illness care model[71-75] to guide network formation and improvement activities. Sites received formal QI training, developed and shared evidence-based changes to support good chronic illness care, including pre-visit assessment, creation of chronic care registries, population management review, a growth and nutrition algorithm, "model" care guidelines, and a guideline to standardize physician disease severity assessment. Communication is multi-modal and involves transparent sharing of performance data.

Data about the patient, disease status, and the care provided are collected during each encounter. Currently, data entry occurs via web-based data capture into an electronic FDA-validated registry hosted by Clinicpace Worldwide. In a recent evaluation of the completeness of data for the 11 variables needed to compute a Crohn's Disease activity score[76] in the existing IBD registry, 8 of 11 had at least 80% completeness, and all fields had >90% completeness for a shorter version of the index.[15] This suggests that the network has

achieved a reasonable level of data quality, although opportunity exists for improvement. At the end of each month, data from the current and previous months are extracted and analyzed to create three reports that are distributed via a password-protected web site: (1) measures of process and outcomes (See Appendix 4 for list of process and outcome measures); (2) remission rates; and, (3) detailed patient information necessary for population-based care management and patient tracking. Each report contains information about the performance of the individual site and summary information of all sites combined. All performance data are shared transparently, enabling sites to identify those achieving unusually good performance so they can learn from one another.

The network has conducted research on the variations in IBD clinical management[14] and development of outcome measures for CD patients.[15] A recent analysis[77] demonstrated substantial improvement in the process and outcomes of care. Process and outcome measures were compared between July 2007 and March 2009 for 1,014 patients (54% male; 72% Caucasian; 62% $\leq$15 yrs). Changes were observed in: (1) an assessment "bundle" (% of visits during which an assessment of growth, nutrition, disease phenotype and disease activity was documented increased from 21% to 86%, $p<0.01$); (2) appropriate dosing of thiopurine increased from 44% to 63%; (3) satisfactory growth status (>10th percentile) increased from 82% to 92%; (4) satisfactory nutrition status (>10th percentile) increased from 76% to 86%; and, (5) remission rate increased from 49% to 64%. Thus, ImproveCareNow has demonstrated the capacity to remove unnecessary practice variation, improve (dramatically) patient outcomes, and conduct research. Addition of CE and QI research components to the work of ImproveCareNow is a logical next step in the maturation of the network.

B4 Advancing QI and CE Research Methods

Participating investigators have developed and tested numerous QI interventions in outpatient settings. For diabetes, Dr. DeWalt and colleagues developed a diabetes disease management program as part of their clinical practice.[78-80] They engaged in a multidisciplinary team, consisting of clinical pharmacist practitioners, diabetes educators, nurses, and care assistants, to help advance the diabetes care, which was primarily driven by individual doctor-patient encounters. This model of clinical care included the use of a registry, the development of care algorithms for all aspects of diabetes care, including insulin and oral anti-diabetic medications, blood pressure medicine titration, cholesterol medicine titration, and subspecialty referrals. After pilot testing this program, Dr. DeWalt and his colleagues performed a randomized clinical trial of 217 patients with type-two diabetes and poor gylcemic control. Intervention patients, who received intensive management from clinical pharmacists and a diabetes care coordinator, had significantly greater improvement in systolic blood pressure (-9 mm Hg; 95% CI -16 to -3) and A(1C) level (-0.8%; 95% CI -1.7 to 0) than control. At 12 months, aspirin use was 91% in the intervention groups versus 58% among controls ($P<0.0001$). Of particular interest, this program was designed to meet the needs of the most vulnerable patients, including those with low literacy, and the outcomes demonstrated more benefit online to those who are more vulnerable.[78] In a similar vein, Dr. DeWalt developed and tested a heart failure self-management education program for patients with limited literacy.[81] This program was also tested in a small randomized clinical trial[82] and is now undergoing testing in a multisite randomized clinical trial funded by NHLBI (5R01HL081257). This novel program helped train patients to manage their heart failure more successfully (patients in the intervention group had a lower rate of hospitalization or death (crude incidence rate ratio (IRR) = 0.69; CI 0.4, 1.2; adjusted IRR = 0.53; CI 0.32, 0.89) and to have more productive interactions with their healthcare teams. Organizations around the world have sought to implement aspects of Dr. DeWalt's program, including researchers and China, Spain, New Zealand, and around the United States.

This project will advance pediatric CE and QI research (and respond to the RFA) by adding a patient-reported outcome data collection system to the informatics infrastructure. This is made possible by Co-I Forrest's work as a PI (U01 AR057956) and Chair of the Executive Committee for the NIH-PROMIS network (which Dr. DeWalt is also a PI), as well as his long history leading federally funded pediatric outcome assessment projects.[83-92] As part of its roadmap initiative,[93] the NIH formed the Patient-Reported Outcomes Measurement Information System (PROMIS) network to develop methods for efficiently and validly examining health outcomes from the perspective of patients.[94, 95] The goal of PROMIS is to develop the components of a novel assessment tool set (using item response theory and computerized adaptive testing) to measure patient-reported outcomes (PROs) regarding physical, emotional and social health across a wide variety of chronic diseases. In addition to a conceptual framework of health and related calibrated item banks, PROMIS has created an Assessment Center and a system for collecting PROs in clinical research that will be used in the proposed study.

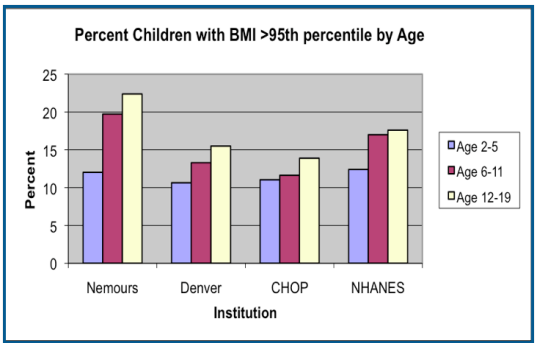
Dr. Joffe has developed history-adjusted marginal structural models,[96, 97] which are a variant of marginal structural models useful in identifying optimal treatment strategies and in allowing modification of the effect of a treatment by time-varying covariates. Dr. Joffe used this approach in estimating the effect of intravenous iron on mortality among hemodialysis patients,[98] and in estimating the effect of hemoglobin variability on mortality in the same population.[96] We propose to advance these innovative statistical methods for CE research by using them to estimate the effects of alternative treatment strategies, with an emphasis on timing of the introduction of biologics, on the outcomes of children with Crohn's Disease.

B5 The Pediatric EHR Data Sharing Network: PEDSNet

As part of its learning healthcare initiative,[1] the IOM stimulated development of Electronic Health Record Innovation Collaboratives, one of which is the Pediatric EHR Data Sharing Network (PEDSNet)—see Letter of Support from Dr. Michael McGinnis. PEDSNet was launched in late 2009 after two meetings held at the IOM. It began with 18 children's hospitals; 7 will be participating in the proposed project. PEDSNet sites are devoted to using their EHR data for research, practice improvement, and transforming outcomes for children with complex conditions. The vision of the network is *to create a national network of pediatric organizations that share EHR and related clinical data to transform children's healthcare.* As a demonstration of their commitment to the successful development of PEDSNet, the Chief Executive Officers from institutions participating in this project have provided letters of support--see Letters of Support.

PEDSNet was created to provide an enabling platform for open-science pediatric CE research and practice improvement projects. This infrastructure will be leveraged by groups of sub-specialists (e.g., pediatric gastroenterologists) organized into network-based collaboratives (e.g., ImproveCareNow network), research groups interested research analyses, and institutions interested in improving care. Thus, PEDSNet is a *network of networks* and, over time, the core group of Children's Hospitals and ImproveCareNow will be augmented by other sub-specialty collaboratives and pediatric healthcare organizations. This sort of organic growth suggests the need for a flexible, federated network, both in terms of data distribution and governance.

PEDSNet Institutions and Data Sharing: As a demonstration of the institution's ability to share EHR data, the PEDSNet consortium initiated a pilot project on exchanging EHR data for analyses related to pediatric obesity. The protocol uses demographic, anthropometric, and diagnosis information from existing visits to support a retrospective review of current standards of care for obesity in children across multiple institutions. In addition to answering an important clinical question, the pilot also highlights issues of data quality and completeness in pre-existing EHR data, as well as logistic factors involved in extracting and transforming such data for use in clinical research. As the first step in validation of exchange methodology, EHR data (2007-2008) were accrued from three pilot institutions for 2.5M encounters involving 630K unique patients (**Figure).** The EHR data correlate well with the NHANES sample cohorts in 2003-2006, although there is variation across sites. These results demonstrate the power of PEDSNet to generate very large clinical samples, the overall utility of cohort construction from EHR data for research, and valuable support for conclusions drawn from the small NHANES cohort, while revealing potential regional variations in patterns of childhood obesity. The project is currently expanding to include assessments of blood pressure and additional centers.



C RESEARCH DESIGN AND METHODS

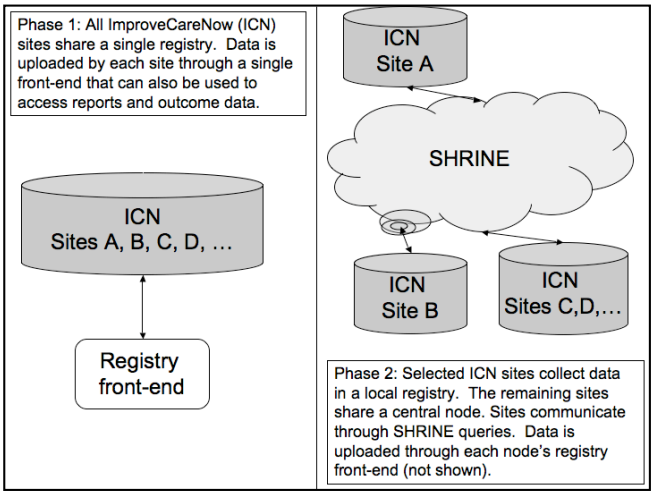
C1 Aim 1: Enhance an existing registry to support a learning healthcare system for pediatric inflammatory bowel disease by capturing needed data directly from electronic health records, improving the quality of collected data using new tools we have developed for recording clinical information during a patient encounter, and facilitating interventions to improve the quality of care for children.

C1.1 Overview: Design of an Enhanced Registry to Address Limitations of Existing Registries. As is typical of many current registries, ImproveCareNow collects data about the patient, the disease status,

and the care provided during each encounter. Data entry occurs via web-based data capture into a central electronic FDA-validated registry hosted by Clinicpace Worldwide in Research Triangle, NC. At the end of each month, data are analyzed to create three reports that are distributed via a password-protected web site. Our proposal seeks to address major limitations of the present registry: (1) data will be captured once during the routine clinical care of the patient so that manual double data entry is not necessary, and (2) data will be housed locally in federated data warehouses rather than in a single central registry with on-demand patient and population management reports. The enhanced registry is built on a modular, open-source, framework. Interoperable modules permit the assembly of specialized registries with little cost and technical effort. If data standardization is in place at the point of care, as is the case within the ImproveCareNow network, the creation of a second disease registry requires only the creation of a new data dictionary. In addition, concepts and variables can be shared, promoting reuse and allowing for cross-registry queries (where permissible). The technical architecture supports the creation of a "distributed" registry. All of the data are housed locally at each site and information is exchanged using distributed queries (creating a "virtual" repository). If the participating sites choose not to participate as a node on the network, they can send their data to a central node, as in the traditional model. In this manner, the technical architecture supports distributed, partially distributed and centralized models. Details of the Technical Architecture can by found in Appendix 3.

Data can enter the system in multiple ways, allowing institutions of various means and sophistication to participate. Sites can feed data directly from their EHR, they can upload files generated by their enterprise reporting team, and for sites without EHRs, they can enter data directly through a web browser. Because the registry is based on the i2b2 framework, a site can also use its registry instance as institutional research data warehouse to meet other institutional data needs (though this is not required). Another benefit to using an i2b2-based registry is that users can take advantage of new features and capabilities that are being contributed by a very active academic users group. The approach of implementing a local repository for both local and distributed query needs can potentially reduce the number of systems that must be supported. By sharing data definitions, queries across registries or specialties are supported. A site can participate in a distributed network only if they agree to use the standard data definitions. The system can support a range of users, from those who simply want to download "canned" reports, to users who want to run ad hoc reports and perform basic data analysis, and via data export, users who want to perform heavy statistical processing and data mining. A site never has to send its data to a centralized node, meaning they are always in control of their data. If a data-owner chooses to leave the registry, it is left with a research database that can function as a local registry. A single node in the network has all the functionality of the larger network, just a smaller pool of patients. The platform has built-in data validation, meaning it is not necessary to extract the data for verification. Users are notified of data errors and failed validation checks. Because it may not be desirable to load data that has not gone through a QA process, data can be "held" in the local registry database and not exposed to the network until it undergoes a cleaning process.

C1.2 Technology Deployment Strategy and Plan. The ImproveCareNow sites will implement the distributed registry through a phased plan (**Figure**). During Phase 1, sites will use the enhanced registry front-end to enter/upload data directly from their EHR into a central ImproveCareNow registry node. Sites will be responsible for extracting the relevant data elements from their existing EHR systems to create the data-upload files. In this phase, all data are centrally housed, and all reports and data extracts for research are obtained from this registry.



In Phase 2, six sites that are Children's Hospitals will implement a local instance of the enhanced registry front-end. These sites will migrate from importing data into the ICN central registry to importing data from their EHRs into a locally controlled ICN registry. The SHRINE technology will be implemented to allow for distributed queries across these networked registries. The remaining sites will continue to participate by entering their data into the central ICN registry. Lastly, three of these six sites will directly extract, transform, and load (ETL) data from their EHRs to their local ImproveCareNow (ICN) registry.

This phased approach allows for a more gradual adoption process, where new functionality is deployed as a step-wise series of small changes. This, in turn, makes it possible for data collection to continue without disruption. We have contracted with Recombinant Data Corporation to assist with the creation of the local registry nodes and the testing and configuration of the SHRINE distributed query network. They are intimately involved in the CICTR SHRINE implementation (see Letter of Support from Dr. Kohane) and have extensive knowledge of the underlying technologies. In summary, our approach will allow for three ways to get EHR data into the registry: (1) export of data from a site's EHR and then file upload to the central ICN registry; (2) export file from site's EHR uploaded to local ICN registry; and, (3) ETL processes from a local EHR to a local ICN registry.

While the local registries are being deployed specifically to house the data from ImproveCareNow, the local registry can be built upon to serve as an institutional data warehouse, if the institution so desires. In this way, the architecture allows for a site to start with a centrally hosted registry and later migrate to a local warehouse, as experience with the tools and technologies grow. Over time, given the strong advantages of distributed, locally-controlled research warehouses, we expect most participants to adopt the distributed model.[30] The local data governance provided by this model alleviates several legal and regulatory concerns that arise from centralized data repositories.[30] However, the proposed architecture does not *require* that decision as a pre-condition of participation; sites may focus their resources in the manner that best suits them.

C1.3 Informatics and Quality Improvement Approaches to Improve Data Quality. We will apply quality improvement methods to enhance uptake of methods necessary to achieve high levels of data quality. Lastly, we will conduct a series of statistical analyses that will be done to identify the best methods for achieving high levels of data quality in an EHR-based registry.

EHR-based Data Collection: We will draw on the ImproveCareNow QI capacity to support sites' application of reliability principles[99] during their conversion to EHR-based data collection. Designing processes for reliability involves three steps: (1) preventing errors by standardizing the process to achieve 80% reliability; (2) identifying steps in the process where failures are occurring, and designing processes to mitigate them; and, (3) learning from the failures to re-design the process using reliability methods.[100] We will make available to the 15 sites that use EpicCare a variety of options for structuring data collection using templates and flow sheets. We will also confirm that the new EHR-based data capture tools meet requirements for variable specifications so that they can be included. Making these options available in their entirety will accelerate each site's IT conversion of their documentation tools and processes to capture high quality IBD data. For non-Epic users, we will work closely with sites by sharing the Epic materials as a starting point and collaborating to ensure that all the necessary data fields can be captured and retrieved from the EHR's data warehouse. Because this conversion will not occur at the same pace across all sites, we have structured our informatics platform to allow a user to continue entering data into the central registry through a web-based portal.

Data Quality Assessment: The ImproveCareNow network currently provides monthly data quality reports. However, the transition to EHR-based data capture during standard clinical care will introduce new challenges. Data will be incomplete unless clinicians enter all necessary data elements reliably. We will take advantage of the modular registry architecture to improve data quality processes and the ImproveCareNow network's existing QI capability to ensure that sites' use the enhanced data quality procedures and to increase the reliability of data collection during sites' transitions to EHR-based data collection. The registry data network will include a data quality assessment process for each file submission at the time of upload to determine if a given data item contains a value that meets the field specification and makes sense in relation to itself and other related data items. In addition to these data management tasks, the registry reporting module will be programmed to automate the process of generating aggregate and site-specific data quality reports that will be distributed back to each site as weekly run charts for use in data quality improvement (see below) and reviewed by the network leaders on a monthly basis. Automation obviates the need for statistical programmers to produce the reports and enables sites to auto-generate reports. Reports about data errors will inform the design of implementation and continued training for sites. Training will be offered by conference calls and during routinely scheduled ImproveCareNow learning sessions.

Statistical Analyses on Data Quality: The change over time in the data entry methods and data quality improvement processes implemented by each site during the 3-year study period provides the unique opportunity to learn about the most effective and efficient ways to enhance data quality. At each site, we will record the time to implement, milestones, challenges and obstacles (e.g., competing priorities among institutional IT departments that delay implementation, difficulties implementing enterprise reporting to output

necessary data, lack of uptake of QI methods), and successes in overcoming them. Each site and the staff from the informatics core will record observations. We will perform a set of quantitative analyses (see below) that examine data quality as a function of time, interventions, and a variety of patient, physician, and practice characteristics. When combined, these analyses will provide a mixed methods (qualitative/quantitative) view on key factors affecting the quality of data collected from EHRs and steps to achieve high quality data.

Data quality is a function of *completeness* and *accuracy*. It is important that data are present for all variables, visits, and patients. It is also important that the data items are valid entries. According to measurement theory,[101] the true value of a variable can be thought of as the sum of its observed value, systematic measurement error, and random error, which can be expressed as: $T = O + E^S + E^R$. Data that are incomplete produce an undefined (missing) value for O. Our data reports, informatics, and QI methods are designed to minimize bias associated with both types of error. Systematic error may occur, for example, if a data element does not have an operational definition that is commonly agreed upon, accepted, or used; the common pediatric health data ontology, which we will develop, minimizes this problem. Random error may occur through inaccurate data entry; the transition to EHR-based data entry (data-in once) mitigates this problem. Observations with larger error terms can be detected as outliers during data analysis.

Data completeness, at the population level, will be measured as the proportion of active patients included in each site's data. Active patients will be defined as those with 2+ visits in a single year (to avoid including $2^{nd}/3^{rd}$ opinion consults) or 2+ visits over two years (to ensure that we include patients with mild variants). All active IBD patients in a practice will be identified from practice registration records. Once these patients have been identified, the registration database will be merged with the clinical database and the patients missing in the latter will be identified. The percent of active patients missing from the database will be the quality measure. Similarly, it will be important to assure that all visits for a patient are included (patient level). This quality measure can be obtained from practice registration records. And, *data item level* quality will be measured as the number of items recorded divided by the number of data values that should be present.

For *Data Accuracy*, two measures will be calculated: (1) whether an item passes range checks and valid value checks and (2) whether an item passes cross-item logic checks (e.g., sequential height measurements should be the same or the second should be higher).

Statistical process control charts will be created for each of the data quality measures and shared with sites. The time points for the significant changes—expanded (more granular) data quality reporting, conversion to EHR data entry, and enhanced QI methods--will be indicated on the charts. The charts will provide a visual display of the effects of each change, and will be a useful feedback tool for practices and physicians.

To create a *summary metric of data quality*, we will score each data field as a 0 (not complete), 5 (complete but not accurate), or 10 (complete and accurate) and divide the sum of these scores across all data fields by the number of data fields; this will yield a mean data quality metric for the visit. The measure could also be summed for a single data field across all visits. *Visit-level and data-field data quality* will be explored to ascertain distribution and variability across practices, physicians, time, and other key predictors.

We will implement *hierarchical linear models with robust estimators*[102] to evaluate visit-level data quality as a function of time and data quality improvement "events"—addition of an expanded set of data quality reports, EHR-based data collection, and enhanced QI. Time will be modeled as a continuous variable with an autoregressive error structure. We will explore time interactions with data quality events to determine if these intervention effects are instantaneous or more gradually accrued over time. The models will additionally include random effects for each practice and fixed effects of likely correlates of data quality such as existence and training of dedicated research staff at a practice, existence and percent effort of local QI support staff, number of clinicians in the practice, and practice sample size. Findings from this work will help to determine the impact of various approaches to enhancing data quality. One key question we will address is whether intense QI support is necessary to obtain research-grade data from EHRs. These findings will have important implications for scaling this method to other conditions and networks.

A second method we will employ will use statistical models conventionally applied to interrupted time series study designs.[103, 104] In these analyses the time trend in visit-level data quality is "interrupted" by implementation of QI methods and training, expanded reporting, or change in data collection type. An autoregressive integrated moving average [(ARIMA(p,d,q)] model will be fit to the data, where p, d, and q refer to the autoregressive, integrated, and moving average parts of the model. Calculations of the autocorrelation function and partial autocorrelation function will be used to determine the values for p, d, and q. Once the model has been identified with respect to values of p, d, q, it will be used for estimation.

To determine if the change in procedures from one phase to the next impacted the data quality, an

impact assessment will be performed.[103, 104] Once the ARIMA (p,d,q) model has been identified, an intervention component will be added to the model and estimated for each of the three types of interventions. Hypothesis tests will be conducted to determine if either component is significantly different than 0 at the 5% two-sided level of significance. If an intervention component is significant, then we will conclude that the introduction of the process changes had an impact on the quality of the data. The nature of the impact will be assessed via visual examination of the time series plots. Appropriate transformations of the measure will be conducted in order to satisfy the assumptions of the statistical testing.

C2 Aim 2: Use Quality Improvement methods to implement enhanced IBD-registry features to enable management of IBD care center patient populations and to increase patient participation in care.

C2.1 QI Project Overview. The central focus of the chronic illness care model is fostering productive interactions between informed, activated patients and their families and prepared, proactive practice teams.[105-111] During the design phase of the ImproveCareNow network, we identified five evidence-based determinants (which we term "key drivers") of health outcomes for patients with IBD: 1) accurate diagnosis and disease classification, 2) appropriate medication selection, 3) appropriate medication dosage, 4) adequate nutritional intake, and 5) appropriate monitoring of growth. Interventions reflecting prepared, activated care teams that could achieve the key drivers and that were conceptually based on the Wagner chronic care model were also identified.

As noted in the AHRQ Registries User Guide,[28] there are significant opportunities to utilize registry functionalities to improve care delivery at the patient and population levels, including: *providing real-time feedback with decision support* (evidence/guidelines), *generating patient-level reports and reminders* (longitudinal reports, care gaps, summary lists/plans, health status), *sending relevant notifications and education to providers and patients* (care gaps, prevention support, self-management), *identifying patients/subgroups* for proactive care, and *providing population measurement and QI reports*. Although effective and efficient use of an enhanced registry has the potential to facilitate several of the ImproveCareNow interventions (population management, patient/family activation or engagement as a component of self-management support, and pre-visit planning) and improve the process and outcomes of care,[112-114] these activities are broad and complex and, therefore, difficult to consistently and reliably implement in practice.

In this Aim, we will leverage the enhanced registry functionalities (that are part of the i2b2 workbench and our Tool Kit, described in Sections B1 and the Technical Appendix) to support advanced QI methods that comprise refining, adapting, and implementing population management processes within each IBD care center, enabling centers to customize longitudinal care for patients as their disease evolves over time. We will also work with a diverse subset of IBD care centers to develop and test patient and family-focused and registry-driven reminders, checklists and prompts aimed at engaging patients and their families to participate more actively in their own care. We will pay special attention to developing the intervention to activate youth and their families, as medication adherence[115] and patient outcomes[116] for chronic disease are superior when both are engaged in the patient's healthcare.

This aim will occur in two phases, a design-development phase and an optimization-implementation phase. In the <u>design-development phase</u>, we will create, pilot-test, and refine enhanced registry functionalities using a structured, iterative, process for new service development common in the advanced application of QI methods in other industries.[68] Qualitative user-centered design methods will be employed to develop the patient-focused, registry-driven tools to enhance patient-focused engagement. Careful development and pilot testing is especially important in a multi-site network, because it allows new approaches to be refined on a smaller scale prior to spreading them.

In the <u>optimization-implementation phase</u>, the network will adapt the new population and patient-level tools using repeated Plan-Do-Study-Act (PDSA) cycles[117] over time. The existing ImproveCareNow quality measures (Appendix 4) and patient-reported outcome measures will be used to determine if changes are resulting in improvements.

C2.2a Design-Development Phase: Testing and refining the enhanced registry and care processes to improve population management. We use the term population management to refer to an integrated set of care processes aimed at improving health outcomes for a population of patients. These processes include: (1) *defining the population* and populating the registry; (2) *dividing the population into categories by defining strata according to a patient's disease activity level* to plan and guide care delivery; (3)

standardizing care delivery within each stratum to create a tiered approach to chronic disease management so that care can be modified if a patient's level of health risk changes over time. A tiered approach requires care teams to develop new roles, ensure coordination of services and needed follow-up using QI methods; (4) *regular population and patient review* to evaluate health outcomes over time and ensure that patients have received the care intended; and, (5) *use of QI reports* to enable sites to track process and outcome measures at a population level over time and compare themselves with others.

Using the QI concept of testing new services under a wide range of conditions, we will select IBD centers representing relatively extreme differences in organizational characteristics and resources to participate in developing and testing registry-supported processes for population management and patient-focused activation. Participation in these activities will be restricted to the IBD care centers that have achieved 80% reliability on an all-or-nothing measure of disease classification and growth monitoring (i.e., classification of disease severity, phenotype, growth, nutritional status and measurement of BMI) and have >90% of their patients in the existing centralized registry. Centers that have not yet achieved high levels of process reliability will be excluded because the complexity of population management processes makes it difficult to implement them until centers have achieved a reasonable degree of care standardization and team re-design. To date, 9 of the 14 centers that have participated for more than 1 year in the disease registry have achieved this level of process reliability. Based on the substantial evidence that a "lead user" approach to developing innovations[118] anticipates the needs of the rest of the market, we will select 5 centers (listed below) that have characteristics of lead users (i.e., they have already anticipated this need and begun to work on new population management approaches). These centers also represent the diverse organizational contexts in which enhanced registry processes will need to operate and, therefore, will provide a wide spectrum of challenges that will need to be overcome to take new approaches to full scale across the ImproveCareNow network.

Site	EHR-Type	Institution	Setting	# of physicians	% Medicaid
UVM	Epic	Academic	Rural	Small	23%
UNC	Home grown	Academic	Rural	Medium	11%
Las Vegas	Intergy	Private	Suburban	Small	12%
Inova Fairfax	Centricity	Private	Suburban	Medium	4%
Nationwide	Epic	Academic	Urban	Large	18%

Intervention Components: Registry enhancements to support the following population management processes will be developed and tested during this phase.

Defining the population and populating and confirming completeness of the registry. As described in Section C1.3 (improving data quality), the network will use QI methods to implement and document the impact of data quality procedures.

Defining strata within the patient population. ImproveCareNow sites already classify patients' level of disease activity, growth (normal, at risk, in failure), and nutrition (normal, at risk, in failure) at each encounter. We will use the measures of disease activity (described in Aim 3) to categorize patients into 3 disease activity tiers (tier 1: inactive-no other risks; tier 2: mild or inactive with growth/nutritional risk; tier 3: moderate-severe, or mild with nutrition/growth failure). We will use these tiers to customize the intensity of follow-up and care management. These disease activity strata are conceptually similar to those commonly used to guide treatment for pediatric asthma.[119]

Standardizing care for each stratum. In order to implement management according to care strata, each IBD care center will have to develop an approach to matching necessary services, defined according to strata classification, to the center's resources. For example, the availability of services, such as psychological counseling or nutritional support, varies across IBD care centers. Some centers have such services available within the practice (typically larger centers), while others refer patients for these services (typically smaller centers). The 5 IBD Centers participating in this phase will use PDSA cycles to test the feasibility of applying the stratification categories. They will work with a sub-set of their patients in order to identify challenges that will need to be overcome to achieve practice-wide implementation. For example, centers will need to design and test processes to provide needed education to patients, ensure follow-up visits are scheduled appropriately, and ways to ensure effective handoffs with ancillary services (e.g., social work) that may not be available during visits. Testing how to implement these processes on a small scale will make the detailed, "how to" knowledge about needed care processes explicit and thus more transferable to other centers. This

will result in tools and experience that can be scaled up and shared with the rest of the network.

Support for pre-visit planning. Pilot sites will work with the informatics team to design and test enhanced registry reports for sites that cannot accomplish pre-visit planning with their EHR, including: *visit-level worksheets pre-populated with registry data to flag patients* with active disease at previous visits, patients meeting criteria for nutrition and growth risk/failure based on previous visits, patients with persistent steroid use (greater than 3 months in duration at any dose), patients hospitalized in the last 3 months, patients with medication doses out of range; and *prompts for health maintenance care*, including tests for medication monitoring (CBC, liver function, influenza vaccines, annual PPD for patients on anti-TNF medications.

Monthly population review. The participating IBD care centers are currently conducting monthly population reviews using a population management report delivered each month as an MS Excel database. Future functionalities will be tested by re-programming the existing tool to make it easier for care center teams to review their population by automatically identifying patients in each strata *not meeting pre-specified outcome and process targets* (i.e., inadequate growth, inadequate medication dosage, such as out of range thiopurine or infliximab dose/kg, sub-optimal drug levels, such as low 6-TG, antibody or infliximab trough) levels, persistent use of steroids), *patients not seen in a pre-specified time interval*, and *patients needing preventive care* (e.g., patients not yet immunized for flu during flu season, patients on thiopurine/biologic who had no recorded WBC or ALT in 3 months). Re-programming the existing tool will create a prototype that will enable test sites to identify any additional requirements for the enhanced registry module that will ultimately allow sites to undertake such a review at any point in time.

QI reporting. The informatics team will also work iteratively with these sites to review current reports and identify specifications for new QI reports and data presentations that will make it easier for centers to compare themselves to one another and to enable “drill-down” analysis to enable care teams to examine data at the individual level that comprise each data point.

C2.2b Designing and testing patient/family engagement tools and reports: During the design-development phase, we will also undertake qualitative work with patients and their families to design registry-generated, patient-focused tools and reports to promote patient engagement in asking questions and ensuring high-quality care. Interviews will explore: (1) knowledge about ImproveCareNow care guidelines for IBD and what families should do to ensure good home-based care, (2) barriers to asking questions about needed care delivery, and (3) how the IBD care center can more actively engage patients and families during visits. Lisa Opippari, PhD, CCHMC Center for Adherence, Maria Britto, MD, MPH, CCHMC Chronic Care Innovation Lab, an expert in user centered design, and Darren DeWalt, MD, MPH, UNC Division of General Internal Medicine, an expert in health literacy, will guide all stages of development of the qualitative study including: designing interview guides, training interviewers, conducting interviews and focus groups, managing and preparing qualitative data for analysis, synthesizing and interpreting qualitative analyses of research data. This group will work iteratively with application programmers to develop the enhanced registry tools.

Participants. We will conduct in-depth semi-structured interviews of IBD patients attending clinic. This study will draw its participants from the CCHMC and Nationwide IBD care centers because of their geographic proximity. Patients will be generally representative of the population in each practice by race and socioeconomic status. In order to maximize the amount of information obtained on facilitators and barriers to active participation in ensuring care according to guidelines, half of the patients will have inactive or mild disease and half of the patients will have moderate to severe disease. As a qualitative study, our sample size is approximate; that is, we anticipate interviewing approximately 30 children and 30 parents per practice (~60 persons), until informational saturation is reached.[120] This will allow for enough individuals to represent a diversity of disease activity and socioeconomic backgrounds. Patients will be identified in multiple ways, with input from participating practices, including identifying patients with IBD diagnoses codes, sending letters from their provider, waiting room flyers, direct provider referral, and patients presenting for IBD visits.

Interviews. Interviews will last approximately 45 minutes and will be audio-taped and transcribed. Respondents will not be identified in any of the files. Each transcribed file will be given a unique 6-digit electronic name. Members of the research team will review each file for accuracy and completeness. A qualitative data analyses software program, ATLAS.ti, will facilitate the analysis. Data from all of the qualitative instruments will be combined and synthesized into one shared database. A deductive and inductive process will be used to create a coding scheme.[121] Prior to data collection, a provisional list of codes will be created from the list of research questions, hypotheses and key variables. Then, we will employ inductive coding techniques as described by Strauss and Corbin.[122] The initial data will be collected, written up, and reviewed

line by line, typically by a participant's response. Beside the responses, categories or labels will be generated and a list of themes will be developed. The themes will be reviewed and then slightly more abstract categories will be attributed to several responses or observations. The responses will then be put into designated qualitative categories. After each interview is coded, text retrievals on specific codes or combinations of codes will be completed. These retrievals enable content analysis of particular topics, which will be followed by displaying the data in a series of matrices to facilitate identification of similarities and differences in themes by respondents (i.e., patients with mild vs. severe disease; low SES vs. high SES patients). Levels of agreement and saliency of themes will be assessed.

Refining the Patient Activation Measure: We will refine an existing short Patient Activation Measure for adults that will be used as an outcome measure in the next phase of this QI Aim.[123] This is a well-validated measure that will require wording changes (but not changes in the concepts) in order to make it appropriate for use in adolescents and families. We will develop a Youth Patient and Family Engagement measure by starting with the existing measure. The co-investigators who are experts in the area and clinicians (n=5) will identify and elaborate any concepts that are relevant for youth and others that are omitted from this measure. Then, small group interviews (n=4 at a time) will be undertaken with youth patients to validate the conceptual framework and to elicit their experiences with feeling engaged in their health care. A similar process will be conducted for parents. Based on this input, we will produce the revised measure(s), which will be pre-tested with 5 youth and 5 parents in cognitive interviews. This approach to ensuring the content validity of measure is consistent with the rigorous approach developed by PROMIS for patient-reported outcomes.[95]

Design and programming of the patient activation tools: Based on the information collected in qualitative interviews, the qualitative research team will work iteratively in PDSA cycles with care center teams, representatives from the patient and family advisory councils described in Aim 4, and the informatics team to develop the content and format of the patient engagement tools and reports, the "IBD care tracker." This group will select specific variables that families and patients with IBD should know and do that are related to a risk of decline or worsening health (e.g., disease knowledge, medication adherence, needed preventive care, follow-up frequency). For each variable selected, the group will also set targets for monitoring that will appear on the tools. They will use data from the qualitative interviews to refine messages and prompts to elicit patients' concerns and questions, and to determine their confidence in managing disease. They will also refine the appearance of the tracker. The tools will present information as a series of questions, a checklist, current performance and desired outcome targets. Successive versions of the form will be tested in small samples of patients using PDSA cycles in the pilot clinics by measuring the proportion of patients who use the tool to identify questions and correctly interpret prompts regarding needed care. IBD care center teams will also test processes for delivery methods. The informatics team will create the IBD care tracker as enhancement to the registry that can be emailed in PDF format, posted to a patient portal, or mailed prior to each visit. Pilot sites will also use QI methods to develop processes to incorporate questions into clinic flow to determine the method by which patients would like to receive the tool (e.g., mail, email).

C2.3 Optimizing-implementing the enhanced registry using collaborative QI methods. Using the enhanced registry to improve population management: Following development and pilot testing of the registry enhancements and new care processes described above, the refined population management process, supported by the enhanced registry, will be spread to the entire ImproveCareNow network. The network will use its existing collaborative QI methods to support IBD care centers in adapting and implementing the new registry enhancements. Center QI support will be provided through sessions devoted to the new tools and processes at existing learning sessions and webinars (each conducted twice a year), monthly conference calls, a listserv, and an extranet document-sharing site and monthly reports. The design of these sessions will draw on the results of pilot testing, the barriers that were identified, and the solutions that were developed. Peer-to-peer support will be provided by care teams that participated in pilot testing. Routine feedback of process and outcome data using the interrupted time series design of run charts and control charts will be used to measure and inform the ongoing optimization of population management processes by IBD care centers.

The primary outcome of optimization of population management will be measures of process changes related to model IBD treatment described in the following table. These measures are already reported monthly. A composite measure will also be computed as the sum of a score of the measures related to model treatment (including timeliness of follow-up) that have the potential to be improved by registry-enabled alerts and population management support. For the composite measure, each variable will be scored present or absent,

for a maximum score of 6 (patients may not be on every medication). Clinical outcomes (that are already measured by ImproveCareNow) will be regarded as secondary endpoints because we do not know how long it will take care centers to achieve reliable implementation of population management. The temporal association between reliable population management and improved clinical outcomes is also unknown.

Domain	Measure
Model Treatment	% of patients with a documented visit within appropriate interval*
	% of patients treated with infliximab, where the dose of infliximab prescribed is at least the dose recommended by the infliximab protocol*
	% of visits where, when disease activity is moderately or severely active and the patient is being treated with a thiopurine, the 6-TGN level has been measured within the previous 90 days*
	% of patients treated with infliximab (at least 5 mg/kg)*
	% of visits where, when disease activity is moderately or severely active and the patient is being treated with infliximab, the infliximab trough level has been measured within the previous 180 days from infusion date*
	% of patients treated with methotrexate where the dose is at least 10 mg/m2 per week*

*indicates variables to be included in the composite measure

Testing the patient activation tools and reports: Following pilot testing and programming of the patient activation tools, the 5 IBD care centers identified above will implement the IBD care tracker for their entire patient population using repeated QI cycles for large-scale testing and implementation (which concentrate on reliability, as described in section C1.3). This will take place after these centers achieve reliable implementation of the registry-enhanced population management processes. In addition to the outcomes listed above, centers will also track the proportion of patients who recall receiving the tool and the average score across all patients on the revised patient activation measure. Data collection for the measures of patient engagement will take place monthly beginning in year 1 to provide a baseline for comparison using the approach described in the Section C3 that describes collection of patient reported outcomes.

Measurement of contextual factors: Variability in the success of QI initiatives may be related to the organizational context in which an initiative takes place. Context includes factors such as organizational and micro-system leadership, culture, QI team skill and capacity. IBD care center characteristics (including number of FTE physicians and staff, setting, type of EHR) will be collected as part of the physician survey described in Aim 3. We will also measure care center contextual factors using an instrument currently being pilot-tested through support from the Robert Wood Johnson Foundation, a study on which Dr. Margolis is a co-investigator.

C2.4 Analysis. Statistical process control charts with control limits commonly used in QI projects will be used to address the two questions related to this specific aim: 1) What is the impact of registry-enabled support for population management on process and clinical outcomes of IBD care?, and 2) What is the impact of patient-focused activation on the process of care, timeliness of care, clinical outcomes, patient-reported activation? Control charts that account for the clustering of patients within centers for each outcome and process measure will be created and reported monthly using the reporting function of the registry. For continuous measurements, x-bar and s charts will be produced. For discrete variables, p-charts will be produced. The points of intervention will be indicated on the charts to allow visual inspection of the impact of the interventions. Charts will be constructed for each individual center and also across all centers. Monthly data on center performance measures exists from all centers for at least 1 year and will be used to establish baseline performance. Standard methods for control chart analysis will be used to determine whether changes to the system of care delivery have taken place at the level of each individual center and across all centers for the evaluation of the impact on population management, and across the 5 centers testing the patient activation tools.[68] In addition, control charts will be constructed combining centers, as appropriate, to compare the following center characteristics visually: type of institution (academic vs. private), setting (rural vs. urban vs. suburban), size of practice (small, moderate, large). Due to the limited number of centers included, no formal statistical comparisons will be done by organizational characteristic.

Using the approach described in Section C1.3, we will also conduct a qualitative analysis of each of the sites implementing the patient/family-focused IBD tracker. Each site will be treated as an independent case study. We will use the existing ImproveCareNow monthly narrative reporting process that IBD centers use to record changes being tested and implementation challenges and obstacles, as well as successes in overcoming them. The combination of quantitative and qualitative analyses will enable us to understand the

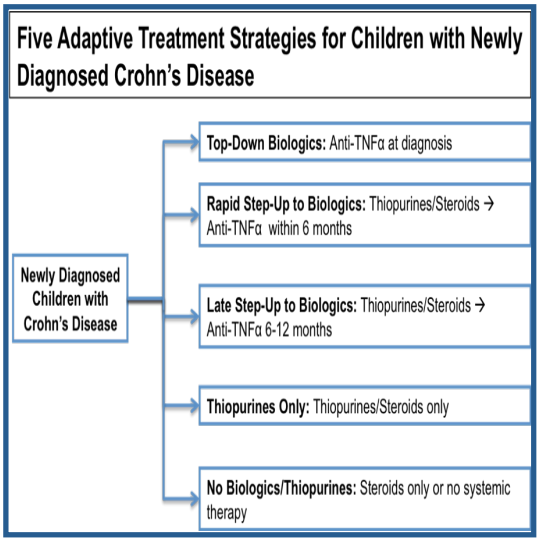
most important components of this intervention, identify the trajectory of learning curves to understand the time to effect, understand different versions of the intervention (including different conditions), and enable us to identify important patient and care center background variables and factors that would degrade performance after implementation.[124] These analyses are also important for future studies.

C3 Aim 3: Use the enhanced registries to compare the effectiveness of alternative treatment strategies for pediatric Crohn's Disease patients, with a special focus on timing of biologic agents

Overview. There are currently several efficacious therapies for the treatment of CD, with the likelihood that others will become commercially available over the next several years. Despite this growing medical armamentarium, the outcomes of patients with CD have not improved in recent years.[54] One reason for this is that optimal strategies for when to initiate and change medications, including corticosteroids, thiopurines, and anti-TNFα biologic agents, are not well-defined, which leads to widespread variation in treatment.[55]

Management of chronic disease is best characterized by treatment strategies--i.e. sequences of decisions that account for prior treatments, patient response, and current health status--rather than isolated treatment decisions. This sequence of clinical decision-making is termed an "adaptive treatment regime or strategy."[56, 125] Conventional efficacy and effectiveness studies evaluate the associations between treatment A versus B (i.e., a single decision) and future outcomes. In everyday practice, however, clinicians make decision according to the changing state of the patient and prior treatment history. They may start with treatment A then step-up to treatment B if response to A is inadequate. The evidence base for comparing two or more strategies virtually never comes from direct comparisons of treatment strategies.

We propose to assess the comparative effectiveness of naturally occurring adaptive treatment strategies among 2,700 children with newly diagnosed CD cared for in ImproveCareNow. Five likely strategies, which reflect both current practice and the controversy surrounding optimal timing for introducing biologic agents into a CD patient's treatment regimen, are shown in the **Figure**. The proposed study will use both non-concurrent (n=1,400) and concurrent (n=1,300) cohort study designs.



Although observational studies can increase the diversity of patients in the sample, as compared with controlled clinical trials, their internal validity is threatened by confounding by treatment indication. We will address this by comprehensively measuring conceptually- and empirically-based factors that affect decision-making regarding biologics. Marginal structural models with inverse probability of treatment weighting will be fit in order to estimate the effects of dynamic treatment strategies on the primary outcome of disease activity and the secondary outcomes of growth, nutritional status, patient-reported health and well-being, and resource use.

Findings from this study will inform: (1) decision-making regarding the best strategy for using anti-TNFα biologic agents in children--one of the IOM's top 25 CE research priorities examined in a priority population; (2) the effectiveness associated with thiopurines and anti-TNFα biologic agents in routine clinical practice; and, (3) the effectiveness of measuring thiopurine methyltransferase genotype or phenotype before initiating treatment with thiopurines. These are the salient knowledge gaps in the management of children with CD.

Study Design: We will conduct two types of cohort studies:[126] a non-concurrent cohort (patients enrolled in the registry before the start of the prospective data collection effort) and a concurrent cohort (patients enrolled after the start of the new data collection). We will also form a combined sample. Because the data collection method and some data elements in the registry will change as a result of the proposed project, it is necessary to analyze the two cohorts separately before deciding on the feasibility of combining them. All patients in the study sample will be newly diagnosed individuals with CD between the ages of 8-17 years at the time of diagnosis. An onset cohort is needed to assess the effectiveness of treatment strategies executed in the first year of disease course, a critical period in the natural history of CD. This approach provides follow-up of patients ranging from 1-6 years, allows for time ordering of medication use, and permits assessment of the dynamic interactions between patients' health status and treatment choices. The main disadvantages of the observational cohort study design are missing data in the follow-up period, attrition (which, according to our

registry data, is 15%/yr due to aging, change in residence, and switching of physician) and confounding by indication. We will address these limitations using innovative analytic methods for observational research, which are described below.

Study Sample: Currently, there are 23 practice sites in the ImproveCareNow network and we anticipate growth to 30 by the start of data collection. Sites are located at universities, children's hospitals, multi-specialty clinics and from a private practice, and are from both rural (Vermont and Maine, for example) and urban settings, producing a study population that approximates the diversity of the US pediatric IBD population. In 2007, ImproveCareNow had approximately 200 newly diagnosed CD patients (15% in any given year are new patients). Since then, the network has experienced growth of 650 total/100 newly diagnosed CD patients per year. We will include 1,400 newly diagnosed CD patients in the non-concurrent cohort, and enroll 1,300 in the concurrent cohort for a total of 2,700 in the combined cohort. Each patient makes, on average, three visits/yr; the combined cohort is anticipated to have 8,100 visits. About 13% of patients are $\leq$10 years of age, 66% between 11-17, and 21% over 17 years (it is common for patients with chronic disease to remain with their pediatric sub-specialist during college and early adulthood); 46% are female and 17% are an ethnic minority.

Data Collection for Patient-Reported Outcomes (PROs): We will use the NIH PROMIS assessment center (www.nihpromis.org) to collect PRO data from parents and children using computerized adaptive testing (CAT) methodology (co-I Forrest chairs the Executive Committee of the PROMIS network). The PROMIS assessment center provides software and hardware that can be used to capture patient-level data for the adaptive tests and related study questionnaires. Our team will develop a PROSPECT-specific web site to collect both patient and parent assessments of health and well-being. This web site will be seamlessly linked to the ImproveCareNow web site, which is where CD patients and their parents will log on. Data from the assessment center are stored in a secure server, data entry is time stamped, and patient name is linked to study identifiers. These data can be downloaded into an Excel spreadsheet, which we will do on a monthly basis, and then imported into the i2b2-distributed network. PROMIS is currently developing LOINC (in collaboration with National Library of Medicine) and SNOMED codes for its items, short forms, and adaptive test scores, and we will store data in the distributed database using these codes. This sets the stage for exchanging PRO data across conditions and studies in the future. Sites will be given three options for obtaining patient and parent reports of child/adolescent perceived health and well-being: patients can be prompted in the pre-visit reminder letter/telephone call to visit the ImproveCareNow web site to complete their health assessment; during the visit, clinics can use kiosks or computers in the reception area; or, discharge instructions will be modified to have the patients/parents complete the PRO data when they arrive home.

Data Collection—Physician Survey: Before we begin the prospective data collection, we will conduct a survey of all physicians in the network to collect data on physician factors that may influence treatment decisions. We will obtain year of fellowship graduation, sub-specialty board certification, gender, practice setting, FTE status, number of CD patients, perceived efficacy of biologic agents and thiopurines, perceptions of risk for both types of medication, and social networking factors that may help explain decision-making.

C3.1. Independent Variable: Treatment Strategy: A patient's current medications are recorded during each patient visit. These data allow us to create longitudinal medication profiles that provide information necessary to form the adaptive treatment strategies. The medications of interest are: anti-TNFα biologic agents (infliximab or adalimumab, which are biologically and clinically equivalent), thiopurines (6-mercaptopurine or its prodrug azathioprine), and corticosteroids (most commonly, prednisone). Our preliminary data suggests a significant level of variability in use of these medications within and between disease severity levels. Within the first two years of patients being diagnosed with CD, we found that 75% received corticosteroids, 65% thiopurines, and 18% anti-TNFα biologic agents (range in use of biologics by practice was 10-40%). We have also observed a recent rise in the past two years in the number of patients receiving anti-TNFα biologic agents, because of the growing belief (and unpublished adult data) that early treatment may alter the long-term course of the disease. These data indicate an adequate level of variability in use of medications to support the proposed comparative effectiveness analyses. Among patients receiving thiopurines, we will assess whether they had a thiopurine methyltransferase genotype or enzymatic level assessed before initiating therapy.

C3.2 Primary Outcome Measure: Disease Activity: This will be measured using the Pediatric Crohn's Disease Activity Index (PCDAI),[76] which includes 11 data elements. The PCDAI integrates gastrointestinal and other systemic symptoms, physical examination findings, laboratory values, and growth

parameters to generate a composite severity score. It correlates well with physicians' global assessment of disease severity, and is the standard measure of treatment response in clinical trials.[127] PCDAI scores >10 (scale 0-100) indicate active disease; ≤10 is considered inactive disease or remission. In some cases, not all 11 components of the PCDAI are collected during the encounter. For visits in which the PCDAI cannot be determined due to missing data (>30% of data elements missing), disease activity will be assessed by a short version, which was recently developed by our group (led by co-I Kappelman) by retaining and re-weighting the components of the PCDAI that are most often reported in medical records.[15] This instrument is strongly associated with the parent index (r=0.86).

Physician Global Assessment: This is a physician's subjective determination of a patient's overall disease activity category (inactive, mild, moderate, or severe) that is made at each visit. We will use this as a secondary measure of disease activity.

Steroid-Free Disease Activity: Corticosteroids are commonly used to treat acute exacerbations of CD.[128] Their utilization is an indication that disease is not in remission, although use of steroid is conflated with physician practice style. Furthermore, corticosteroid dependence is frequent among children and adults with CD[129, 130] and is associated with long-term toxicity. Because of these side effects, one of the key objectives of CD management is to achieve steroid-free low disease activity. Thus, we will create a second outcome measure of disease activity that combines absence of steroids (i.e., steroid-free disease activity) with remission (PCDAI<10) to form a binary outcome measure of steroid-free remission.

C3.3 Secondary Outcome Measures: Anthropometrics, Patient-Reported Outcomes, and Resource Use:

Anthropometrics: Growth failure is one of the hallmark features of CD at diagnosis[131] and an indicator of inadequately managed disease.[132] To assess growth, we will obtain height velocity z-scores. We will use BMI z-scores and weight z-scores as measures of nutritional status.

Patient-Reported Outcomes: We will use the PROMIS pediatric item banks, including new measures that we are currently developing (see Preliminary Studies). All PROMIS PROs are developed using a rigorous set of standards, a combination of qualitative methods (focus groups, literature reviews, and cognitive interviewing), confirmatory factor analysis to ensure construct unidimensionality, and item response theory.[94, 95] We plan to analyze the scores (person's estimated ability level or theta on the latent trait) for physical activity, physical functioning, pain, fatigue, sleep functioning, anger, anxiety, depression, experience of stress, subjective well-being, and family belonging. All item banks will be given as computerized adaptive tests, which require as few as five items to provide each latent trait score. The full set of item bank scores will yield a profile of perceived health and well-being for each child. In addition, we will use the IMPACT III as a fixed-length, short form measure of IBD-specific perceived health.[133] IMPACT III is a minor modification of the IMPACT II.[134, 135] It includes items with visual analogue response options, has an internal consistency of 0.96, and a test-retest reliability of 0.90.[135] The questionnaire provides scores for symptoms (general and bowel), social concerns, body image, and emotions.[136] This instrument has been associated with clinical improvement in trials of IBD medication efficacy.[53]

Resource Use: Inadequately treated CD can result in frequent specialist visits, increasing intensity of medication, gastrointestinal surgery, and hospitalization. These measures have become frequent secondary endpoints in randomized controlled trials and studies of clinical effectiveness.[65, 137] Anti-TNFα treatment is significantly more expensive than standard treatment (approximately $50,000-$60,000/year versus $3,000-$5,000/year). On the other hand, it is possible that biologics are more effective than thiopurines and corticosteroids at reducing total healthcare utilization by lowering the risk of acute exacerbations and the need for urgent visits, hospitalization and surgery. Data from all sites will be used to assess healthcare utilization (number of specialist visits, surgeries, and hospitalizations).

We will also perform a more detailed analysis of healthcare expenditures at the six pediatric sites that install a local instance of the i2b2 database (Children's Hospital of Philadelphia, CCHMC, Nemours, Nationwide Children's Hospital, University of North Carolina, Denver Children's Hospital). These sites account for approximately 50% of the total study sample. CPT-4 procedure code information will be obtained for specialty visits, endoscopies, and surgeries; medications prescribed (NDC codes and dosage) and information on the APR-DRG and length of stay for hospitalizations will also be extracted. To eliminate variation in charges due to differences in fee schedules, we will standardize expenditures to a common fee schedule, a method that we have used previously.[138, 139] For inpatient stays, we will use the 2010 Maryland Health Services and Cost Review Commission data to obtain a per diem cost (APR-DRG case-mix index weight*conversion

factor) and will multiply this by the length of stay to obtain a standardized charge for each hospitalization. For outpatient services, we will standardize charges using the 2009 Medicare fee schedule, and the 2009 Red Book of wholesale drug prices for medications. This standardization methodology will allow us to form a set of measures for total healthcare, outpatient, inpatient, and medication expenditures.

C3.4 Potential Confounders: The most important threat to the proposed study's internal validity is selection bias, which results from confounding by indication. In other words, the same factors that influence outcomes also "indicate" the need for specific treatments. An experimental design balances these confounders across treatment groups. In an observational study, one must attempt to measure and adjust for as many potential confounders as possible to minimize bias associated with parameter estimation. Our approach to elucidating these confounders is two-pronged: conceptual and empirical. On the conceptual level, we will operationalize Iezonni's model of health risk, which comprises co-morbidity, physiological stability, disease severity, disease burden, and perceived health.[140] In addition, we performed a modified Delphi survey of 27 physician leaders in the ImproveCareNow network (response rate = 75%), asking them to identify (Round 1), then prioritize (Round 2), the disease, patient, clinical/physician, and health system factors that influence their decision to use biologic agents among patients with CD. Results are shown in the **Table**.

Factors influencing a clinician's decision to use biologics among children with Crohn's Disease: Results from a Modified Delphi	**Mean Importance Rating (1 low, 10 high)**
Fistulizing/internal penetrating disease	9.5
Disease severity	9.0
Failure of response to immunmodulators/thiopurines	9.0
Perianal disease, including fistula	8.8
Steroid dependence/steroid resistance	7.5
Side effects of other medications	6.9
Growth failure	6.9
Extensive small bowel disease	6.4
Stricturing disease	6.2
Nutritional failure	5.8
Family preference	5.6
Extra-intestinal manifestations	5.5
Lack of adherence to other medications	5.4
Risk of lymphoma (influencing non-use)	5.4
Disease location	5.3
Willingness to undergo infusion	5.1
Cost of biologics to family (influencing non-use)	5.3
Degree of ulceration seen during colonoscopy	5.1
Logistics of giving biologics	4.0
Age of patient	3.7

The primary clusters of factors are disease extent and behavior, disease severity and activity, prior treatment response, and growth/nutritional failure.

Disease Extent and Behavior: At each encounter, clinicians record presence of fistulizing or penetrating disease, perianal manifestations, extra-intestinal manifestations, strictures, and disease location. We will characterize the extent and behavior of CD using the internationally accepted Montreal classification.[141]

Disease Severity: Many of the factors in the Table are indicators of disease severity, which we define as the extent, impact, and stability of the disease. Both the PCDAI and the physician's global assessment of disease activity will serve as an outcome and a predictor in statistical analyses. The C-reactive protein (CRP) will also be examined, as it is the biomarker of choice for routine assessment of intestinal inflammation due to CD. CRP tracks with disease activity,[142] severity, and nutritional status.[143] Lastly, we will use the patient reported outcome measures, described above, as generic measures of disease impact.

C3.5 Statistical Analysis: The Figure (Aim 3, first section) of the five adaptive treatment strategies is an idealized view of sequential treatment decision making among clinicians managing new-onset pediatric CD. Our interest, however, is to examine naturally occurring treatment strategies, some of which are likely to diverge from the idealized view. To account for these variances, which themselves are informative, we will model treatment choice at each visit as a function of past health status covariates and previous treatments. This approach accounts for the dynamic interaction of treatment and outcomes and is most appropriate for longitudinal observational studies. We will organize our analyses into three objectives: (**Obj 2.1**) assess the disease, patient, provider, and health system factors associated with choice of initial treatment and subsequent use of thiopurines and anti-TNFα biologics; (**Obj 2.2**) evaluate the association of each dynamic treatment regime with the primary (PCDAI) and secondary (anthropometrics, patient reported outcomes, and resource utilization)

outcomes, singly as univariate repeated measures outcomes and in combination as correlated repeated measures outcomes; and, (**Obj 2.3**) explore alternative statistical methods for estimating the effectiveness of anti-TNFα biologics and thiopurines on outcomes and for addressing three potential threats to internal validity: confounding by indication, missing data, and attrition.

Power Calculation. For our power calculation,[102] we assume a repeated measures outcome using the PCDAI and, for the entire sample, expect that there will be eight measurements on average. The spacing of each measurement will be approximately four months. Our calculation compares the effectiveness of biologics versus standard treatment, and the 1-year period prevalence of biologics is estimated at 20%. The mean and SD of the outcome measure are approximately 30 +/- 8 for all CD patients.[144] We assume the ICC for repeated measures of the activity index to be 0.6 based on our existing data. Under these assumptions, we calculated the minimum difference in rates of change in outcome between subjects using and not using biologics that can be detected with 80% power using a type I error rate of 5%. The minimum detectable difference was 0.45 units of change in the index score/year for the concurrent cohort and 0.32 units change in score/year for the combined cohort. These numbers are optimistic because of the reduced power associated with inverse probability weighting, but we expect that the effect of biologics is substantially larger than this, suggesting very high levels of power, which is primarily due to the very large study sample we will assemble in this project.

Objective 2.1: Determinants of Use of Biologics and Thiopurines

To better understand variation in treatment decision-making, we will estimate the associations of disease, patient, provider, and health system factors with use of both types of medications using a framework that we have previously developed for clinical decision making.[145] In the first set of analyses, we will model choice of initial therapy (three months from the time of diagnosis) as a multinomial outcome--anti-TNFα, thiopurines and steroids, steroids, and others--using logistic regression for unordered polychotomous outcomes and generalized estimating equations to account for patient clustering within physician.[102] We will also conduct analyses in which each physician is entered into the model as a random effect to determine if we can identify characteristic practice styles among physicians--i.e., some who always use biologics as top-down therapy and others who always start with steroids only, etc. (which is relevant for the instrumental variable analyses discussed below). Next, we will conduct survival analyses using Cox proportional hazards models[125] in which the outcome is time to initiation of treatment, accounting for censoring of patients. In these models, the outcomes of interest will be anti-TNFα and thiopurines and the independent variables will be the same as the multinomial regressions, specifically: severity, activity, behavior, and extent of disease; patient age, sex, race/ethnicity, insurance status, and year of diagnosis; and, provider attitudes about perceived efficacy of anti-TNFα biologics and thiopurines, safety of these medication and steroids, years in practice, gender, and practice setting. We will also enter medications and prior treatment response as time-varying covariates to determine if specific patterns of medication and treatment response are predictive of use of biologics or thiopurines. Choice of therapy will be modeled as a repeated-measures categorical outcome using unordered polychotomous logistic regression, using generalized estimating equations to account for multiple observations within a subject; these models will include prior therapy and other time-varying covariates, including disease activity, as well as baseline covariates. Findings from this objective will enable a comprehensive characterization of the determinants of decision-making regarding use of biologics and thiopurines, which is a necessary prelude to estimating treatment effectiveness.

Objective 2.2: Association of Adaptive Treatment Strategies with Outcomes

In this objective, the analytic goal is to compare different treatment strategies and optimize those strategies--a sequence of treatment decisions that account for a patient's health status trajectory--for minimizing disease activity. Standard multivariable regression adjustment methods are generally inadequate when confounders vary over time, are affected by the independent variables, and influence the outcomes.[146] In these cases, confounders act as intermediate variables in a causal pathway, and conventional approaches can lead to biased parameter estimates.[147] Marginal structural models (MSMs) can accommodate observational study designs that require adjustment for time-varying covariates and outcomes that dynamically interact with one another.[148, 149] The models adjust for confounding by using all the predictive factors, such as health status and prior treatment response, to produce weights that are functions of the inverse of the patients' estimated probability of receiving the treatment they actually received at each time point in the study.[150]

MSMs are the statistical model of choice for estimating the effects of time-varying treatments in the

presence of confounding by variables affected by the treatment and have recently been extended to estimating the effect of "dynamic" treatment strategies on outcomes, although they are highly complex to implement. They are increasingly being used to obtain causal estimates in observational studies,[151] and are just now disseminating to the field of CE research. **The use of MSMs to estimate the effects of adaptive treatment strategies on patient outcomes is an analytic innovation of the proposed project.**

We will consider two approaches to using MSMs for estimating the effect of dynamic treatment strategies. In the first, we will model comparisons of the five adaptive strategies considered above. Let Y_t denote the outcome in a given subject at time *t*. Let *g* denote a treatment regime, and let Y_t^g denote the outcome that would be seen in a given subject at time *t* under regime *g*. We model the expectation of Y_t^g as a function of time and regime using a model like $E(Y_t^g) = \alpha_0 + g\alpha_1 + t\alpha_2 + gt\alpha_3$. This model allows the mean value of the outcome under each regime to vary linearly as a function of time and for the slope to vary for different regimes. The model is easily extended to allow for nonlinearities. The model is estimated using (stabilized) inverse probability of treatment weights, in which each subject's weight at any given time is the inverse of the product of the probabilities of following his observed treatment pattern at each time until that time given prior treatment and covariate histories. Because a given subject may follow a course of treatment that is consistent with more than one regime, the subject may contribute to the estimation of the outcome under more than one regime, and so the variance estimates will be adjusted appropriately for this. Tests and confidence limits for contrasts of regimes are easily obtained from the covariance matrix of the estimated parameters.

The above strategy allows comparison of the alternative regimes and selection of the optimal regime among those five, overall and for sub-classes of patients. We will want to consider comparison of a larger number of regimes and selection of the optimal one in this larger class. For example, we can consider strategies in which initiation of anti-TNFα biologics depends on crossing some disease-activity threshold *h*. We can then model the mean of disease activity under a given threshold *h* over time $E(Y_t^h) = \alpha_0 + f(h)\alpha_1 + t\alpha_2 + f(h)t\alpha_3$, where $f(h)$ is a vector function of the threshold; the model allows flexible modeling of the effect of threshold, and can be extended to allow nonlinearities in time and allow the effect of different thresholds to vary with the initial level of disease activity.

As an alternative strategy, we will also use history-adjusted marginal structural models to estimate the effect of anti-TNFα biologics. These models allow one to model the effect of a static course of treatment from any time *k* as a function of biologic agents, and allow modification of the effect of treatment from that time on by a time-varying covariate $E(Y_t^{\tilde{a}_k} \mid \overline{L}_k, \overline{A}_{k-1}) = \alpha_0 + \alpha_1 L_k + \alpha_2 \tilde{a}_k + \alpha_2 \tilde{a}_k L_k$, where L_k refers to covariates measured at *k*, $\overline{L}_k$ refers to the history of covariates *L* through *k*, $\overline{A}_{k-1}$ is the history of treatment through *k*-1, and $Y_t^{\tilde{a}_k}$ is the outcome one would see at *t* if a subject followed received her or his observed treatment through *k* and treatment plan $\tilde{a}_k$ subsequently. The models allow us to examine how the effect of subsequent treatment is modified by covariates at *k*. This, in turn, allows one to choose the optimal static treatment plan $\tilde{a}_k$ at each time *k*, and, by considering these in sequence, the statically optimal dynamic plan. This approach will complement the approaches outlined above.

Similar modeling strategies may be adopted for other outcomes (e.g., disease-free remission); for categorical outcomes, logistic models may be substituted for the linear ones outlined above. Because steroids mask disease activity, the approaches above to modeling disease activity may be biased. One strategy for dealing with this is to consider, for subjects who are on steroids, the disease activity as a right-censored outcome (i.e., a subject who is on steroids and has a disease activity index of 10 would have had a disease activity index of at least 10 had s/he not been on steroids). Thus, in addition to the linear regression MSMs considered above, we will fit censored regression MSMs analogous to those considered above.

<u>Use of Thiopurine Methyltransferase (TPMT) Testing: Effect Modification</u>

Before starting thiopurines, clinicians may choose to check TPMT genotype or enzyme activity (the two are comparable for decision-making). Knowing the TPMT status allows for more appropriate dosing of thiopurines;[152] too high a dose can have untoward immunosuppressive effects. Clinicians check TPMT for about 60% of patients;[14] whether this has a beneficial effect on patient outcomes is unclear. We will examine the effectiveness of checking TPMT status before initiating thiopurines, treating the variable as an effect modifier in our effectiveness analyses.

Objective 2.3: Alternative Statistical Methods for Estimating Treatment Effectiveness

Multivariate Matching: We will use multivariate matching to analyze the effectiveness of timing of biologics while controlling for confounding by indication, a method that has recently been developed for outcomes research.[153-155] The research questions of interest are effectiveness of biologics or thiopurines. Assume that a patient who receives a biologic agent is a case. Controls are matched to a case by identifying patients who look very similar on all covariates that may affect either treatment decision-making or outcome. Once these groups are formed, we can estimate treatment effectiveness, examining associations of case/control status with disease activity, patient reported outcomes, and resource utilization. Silber and colleagues have used this technique in a study on optimal timing of neonatal infant discharge on costs of care.[154] Once patients are matched, we will use an intention to treat analysis to analyze outcomes for the rest of the observation period.[156]

Instrumental Variables: Another approach to reducing the potential impact of endogeneity (a correlation between medication choice and the error term) is use of an econometric technique called instrumental variables (IV). An "instrument" is a variable that is strongly correlated with the independent variable but not independently associated with outcome. In this study, one potential instrument is physician preference for treatment,[157, 158] which assumes that treatment choice by a physician is unrelated to patient characteristics. This could be the case if a physician nearly always prescribes biologics for newly diagnosed patients or almost never prescribes them (assuming the patient has moderate/severe disease). If we do detect such divergence in preference, we will implement an IV analysis to re-estimate comparative treatment effects.

Attrition. We will conduct analyses of patients continuously in the study population versus those who leave the cohort. It is possible that, even with adjustment for observed baseline differences, attrition bias can lead to inappropriate parameter estimates. One approach we will explore to limiting attrition bias is creation of an instrumental variable.[159] A logit regression model is fit to estimate the probability that a subject leaves the study. Fitted probabilities are used to form an instrumental variable that will be entered into subsequent models as a control variable.

Missing Data: Aim 3 proposes several methods to enhance front-end data entry. We recognize, however, that missing data are inevitable both with EHR-derived data-sets and in longitudinal studies. Data may be missing because of incomplete inclusion of patients or missing data fields. Complete case analysis inevitably introduces bias due to informative missingness and produces "bouncing" data sets if missingness is associated with specific lists of regressors. The simplest approach to missing data is deletion of cases, in which all patients with data necessary for the analysis are excluded. For regression analysis with a large number of covariates, this obviously can reduce power and introduce bias into parameter estimation. We propose preliminary steps of multiple imputation for missing data.[160, 161] Multiple imputation separates the missing data model from the response model, provides a unified method for imputing missing values of outcomes and covariates, allows for a broad use of auxiliary variables to impute missing data, and thus avoids bias arising out of mis-specified models.[162] In addition, multiple imputation modeling permits sensitivity analyses for informative dropout leading to missingness not at random (MNAR). Tools for multiple imputation are available in Stata v. 11 (Stata Corp, College Station,TX; 2009). Variance estimates will be derived using the method described by Rubin.[160]

C4 Aim 4: Develop a governance structures for the network that engages patients and provides oversight of privacy, confidentiality, and data access, as well as scientific and technical concerns

The distributed pediatric data network will serve as a proof-of-concept regarding informatics, analytics, learning, and governance for PEDSNet. To meet the goal of forming a federated governance model based on the principles of data sharing, collaboration and learning, and user-led comparative effectiveness research, we will give substantial attention to the technical, legal and ethical, business, and social barriers that arise as a result of large-scale, multi-institutional data sharing using information from human subjects. In this section, we describe the current state of PEDSNet governance and patient/family participation and a set of methods that we will use to shape a future state that supports the long-term growth and sustainability of the federated network.

C4.1 PEDSNet Governance: As a federation of several pediatric children's hospitals and practice-based research and improvement networks devoted to sharing their EHR data for CE research and

improvement (i.e., a network of networks), PEDSNet must have higher-level governing structures, guidelines, and policies that are in addition to and consonant with the security, privacy, and data sharing policies of the individual organizations.[17] The creation of trust within the network requires that all participants adhere to these policies. PEDSNet was created to serve this role, although we recognize that there are a variety of challenges to achieving our vision. In its early stages, PEDSNet will be in engaged in a small number of proof-of-concept studies, such as the proposed project. Each project will be crafted in collaboration with sub-specialty networks and investigators. As PEDSNet matures, it will transform into a more distributed model of governance in which its primary functions will be to maintain and expand the informatics infrastructure, articulate principles that research projects must adhere to in order to gain access to data, ensure that data are protected, maintain and build a common pediatric data ontology that is useful for pediatric QI and CE research, and serve as a community for sharing of analytic methods. (For example, in the proposed project, we will develop new methods for causal modeling of naturally occurring adaptive treatment strategies, and these methods will be applicable to any chronic condition. PEDSNet will be the forum for exchanging these methods and knowledge, freely and openly, and disseminating new analytic methods developed in the future.)

PEDSNet Governance**:** PEDSNet has developed a set of by-laws that include a committee structure. PEDSNet committees meet by teleconference on a biweekly to monthly frequency. Email exchange and a network collaborative wiki site augment communication. We will continue to hold annual meetings at the Institute of Medicine in which representatives from each institution and network convene to evaluate network progress and discuss strategy. PEDSNet committees are:

(1) Executive Committee (Chair: co-I Forrest [Children's Hospital of Philadelphia])—Composed of a Chair, Vice Chair, Sub-Committee Chairs, and 3 at-large members; empowered to carry out day-to-day operations, manage finances, oversee the work of the sub-committees, and conduct periodic conflict of interest reviews of all PEDSNet scientists.
(2) Science Sub-committee (Chair: co-I Kelleher [Nationwide Children's Hospital])—Includes representation from each participating sub-specialty network and PEDSNet health services and outcomes researchers; obtains input and feedback concerning important aspects of the design of data collection infrastructure to ensure that it supplies necessary information for CE research; reviews research proposals to prioritize and ensure highest quality study designs, execution, and analysis.
(3) Network Sub-Committee (Chair: Consultant Del Beccaro [Seattle Children's Hospital])--Establishes threshold criteria for institutional and sub-specialty network inclusion and exclusion in PEDSNet; oversees template based agreements, which allow each institution to retain its own data while granting to the entire group a shared ownership of the ability to aggregate the data across sites for the purposes of research; nominates candidates for Chair of Executive Committee (who appoints the Vice Chair and Sub-Committee chairs) and at-large members.
(4) Technical Sub-Committee (Chair: co-I Kahn [Denver Children's Hospital])--Oversees compliance and technical standards; manages data governance and pediatric ontologies; oversees development of the informatics architecture for participating institutions; develops simplified procedures for giving researchers, both within and outside member institutions, access to the data for analyses; develops a plan for data sharing that offers a viable and reasonable mechanism to allow qualified researchers access to these databases, oversees security and privacy; and, develops common IRB language to allow use of de-identified data.

Each research activity, such as the proposed project, has a research management group (RMG) that is led by the Principal Investigator. The RMGs are responsible for executing and managing research operations supported by extramural funders have substantial discretion on how to conduct the work, and include liaisons to the Science and Technical sub-committees. The RMG for this project will be responsible for monitoring the performance of Recombinant Data Corporation. Their efforts are found to be lacking, the committee will make recommendations on whether to terminate the contract and identify a suitable alternative.

Patient and Family Councils: Each center in ImproveCareNow has 1 patient and 1 family representative who participate in the QI activities. The input these individuals provide is ad hoc and specific to each center. As part of the proposed project, we will form two national patient and family councils composed of these individuals. The Patient Advisory Council (to be chaired by Jill Plevinsky, a senior at George Washington University and Co-Chair of the Crohn's and Colitis Foundation National Youth Leadership Council and the

Parent Advisory Council (to be chaired by Steve McPhail, CEO of Expression Analysis, Inc (see Letters of Support) will be convened on a bimonthly basis by teleconference (and more frequently as needed), will work collaboratively using the project's wiki space, and a sub-group will participate in the annual governance meetings. The full Councils will provide review and comment on all aspects of the proposed project, will help design the QI interventions, and provide input into the evolution of the governance of PEDSNet. Our goal is to fully engage patients and their families in the learning process.

C4.2 Pediatric Health Data Ontology: The NICHD is working to harmonize pediatric terminologies (http://www.nichd.nih.gov/health/clinicalresearch/terminology/), but those efforts are currently limited to neonatal and infant examinations. It is our goal to build upon the work of the NICHD to create a harmonized ontology that crosses disease and service sector categories in order to promote data sharing and interoperability for a pediatric learning healthcare system. This represents a critical step in transforming both project-specific and standard-of-care EHR data into a practical and useful tool for CE research and practice improvement. While portions of this work will involve conceptually straightforward extensions to existing ontologies (e.g. LOINC, RxNorm) initially created around care of adults, more fundamental development is needed in areas such as growth and development, and diagnostic categories (e.g. congenital disorders) more common in children than adults, to create an ontology that accurately reflects pediatric care. The NICHD's efforts address some of those relating to growth and development, but terminologies relating to congenital disorders and chronic disease are lacking. We intend to fill that void. Our approach to this will be a pragmatic balance of breadth and utility, in that we will produce intermediate ontologies that enable current QI and CE research while continuing to develop a comprehensive system. In the proposed project, we will start by harmonizing a pediatric rheumatology ontology developed for an ARRA-funded GO grant with the one we develop for pediatric inflammatory bowel disease (Consultant Dr. Robert Warren is a PEDSNet leader, and developer of the pediatric rheumatology ontology. He and co-I Charles Bailey will lead our ontology work. Working with Science Commons, PEDSNet will establish policies and processes that encourage investigators to make their measures available to others, so that questions can be reused, reducing the variance that occurs when groups create questions that cover much of the same content, but have a slightly different meaning. We will annotate all of the study data with concept-unique identifiers (CUIs) obtained from the UMLS.[163, 164] When applicable, we will annotate the data so that it conforms with emerging CDISC standards like CDASH and BRIDG (www.cdisc.org).

C4.3. Elucidating Barriers to Data Sharing | Creating Solutions: PEDSNet will create a variety of assets including data, novel approaches for QI, new CE research and other intellectual products, and policies. However, an enormous amount of work will be required to develop the informatics infrastructure and policies that enable data sharing. We are proposing to work with Science Commons to explore how to take advantage of network-based, open source or "peer production" approaches to accelerate PEDSNet's work. Specifically, we will identify and address governance issues that are most important to the PEDSNet community and make information about the solutions we develop publicly available. We will concentrate on five types of issues: (1) intellectual property including ability to copy, distribute and use the data, incentives for knowledge and resource sharing; (2) IT issues such as methods to promote application development, ontology development, metadata, methods for data security, data access; (3) policy issues such as privacy, human subjects, IRB; (4) engagement of institutions and the research community and trust building; and, (5) sustainability of the network. Our strategy is to identify barriers and then develop solutions for PEDSNet. We anticipate that others will have similar interests so we will develop ways to share our solutions globally. This will lead to an adoption phase that may recruit other developers (informatics, ontology), which is how i2b2 grew.

Stakeholder Structured Interviews: In collaboration with Science Commons, we will develop a semi-structured interview that will address the five types of issues described above (IP, informatics, policies, engagement, and sustainability). These questions will be framed within the context of Benkler's model of information production—i.e., how individuals, organizations, or networks produce new information.[165] The interviews will elucidate stakeholder perceptions of current state of information production and desired future state. The key dimensions of this model are: market versus non-market uses of information and IP (e.g., we currently make or want to make money from our data assets v. our data assets are available for the public domain); exclusivity of data—use is restricted to data holders only, limited others in the PEDSNet consortium, open-access with appropriate authorization and authentication, or fully unrestricted open-access; and, organization of information production is individual scientists and teams, ad hoc collaborations, firms, networks,

or the global community. The interviews will be done with leaders of children's hospitals and academic health centers, physician sub-specialty networks, sponsors (such as the NIH, AHRQ, and disease-specific funders), professional societies, and journal editors. (Many stakeholders from these groups have already demonstrated their support for this work—see Letters of Support.) These interviews will be done to capture themes and ideas, while applying little structure on the content of responses. Mode of administration will be via telephone and digitally recorded. Transcripts will be content analyzed using Atlas.ti V6, which is a software package commonly used for qualitative text analysis. Once the issues related to these four domains are enumerated, we will work with Science Commons to identify a variety of potential solutions.

Consensus Development of a Solution Set for PEDSNet Infrastructure Development**:** We will use a modified eDelphi process[166] to elaborate on the preliminary solution set and then prioritize the activities. The eDelphi will be done in 3 rounds using Email- and internet-based surveys to obtain experts' opinions. Experts will be nominated by participating PEDSNet institutions and sub-specialty networks, and will seek involvement from other stakeholder groups such as professional society leaders and journal editors. In the first round, panelists will be given a short synopsis of the identified barriers to development of an open-science, pediatric learning healthcare system. They will then be presented with the solution set we have developed and asked to rate the importance of each solution; they will also be given the opportunity to suggest additional solutions. Revisions to the solution set will be made, and the second round will provide these, along with round one responses. Panelists will be asked to rate the level of importance for each of the revised solution sets. A third round will be done using the same approach as Round 2. This process will generate a consensus-derived solution set for overcoming the barriers to forming a federated network devoted to sharing EHR data for CE research and practice improvement. Future work will be devoted to implementing these solutions.

C5 Project Impact

We will extend our existing disease registry for pediatric IBD to create an enhanced modular registry by: (1) linking the registry to electronic health records; (2) using open-source software to develop new applications that support QI; and, (3) connecting the registry to a distributed pediatric data network using i2b2/SHRINE as a platform for data standardization, storage, and sharing. This project will provide new evidence for optimal timing of biologic agents (an IOM top 25 CE research priority) for children (a priority population) with Crohn's Disease. Our CE research approach is to model naturally occurring sequential patterns of treatment decisions (i.e., dynamic treatment strategies); these methods will advance the field of CE research on chronic disease by creating statistical models for determining the optimal longitudinal treatment approach for patients. Supported by a robust clinical informatics infrastructure, the proposed project will fuse CE research with QI within a network of clinicians who coalesce around a common purpose. This innovation is made possible by this RFA and existing infrastructures (i2b2/informatics, ImproveCareNow, PEDSNet) that we are combining and expanding in unique ways. The new informatics and research infrastructure we develop will demonstrate concretely how to build and leverage for QI and CE research a national distributed pediatric data network—a model that can be scaled to all pediatric sub-specialties and millions of children, ultimately creating a pediatric learning healthcare system, as envisioned by the IOM.

Bibliography and References

1. Olsen L, Aisner D, McGinnis JM. *The Learning Healthcare System.* Washington, DC: National Academy Press;2007.
2. Field D, Sansone SA, Collis A, Booth T, Dukes T, Gregurick SK, Kennedy K, Kolar P, Kolker E, Maxon M, Milliard S, Mugabushaka AM, Perrin N, Remacle JE, Remington K, Rocca-Serra P, Taylor CF, Thorley M, Tiwari B, Wilbanks J. Megascience. 'Omics data sharing. *Science.* 2009;326(5950):234-236.
3. Kaye J, Heeney C, Hawkins N, de Wries J, Boddington P. Data sharing in genomics - re-shaping scientific practice. *Nat Rev Genet.* 2009;10(5):331-335.
4. Margolis P, Halfon N. Innovation networks: a strategy to transform primary health care. *JAMA.* 2009;302(13):1461-1462.
5. Metzger Jane. *Using computerized registries in chronic disease care.* Oakland, CA: California HealthCare Foundation;2004. Available at: http://www.chcf.org/documents/chronicdisease/ComputerizedRegistriesInChronicDisease.pdf.
6. Dougherty Denise, Conway Patrick H. The "3T's" road map to transform US health care: the "how" of high-quality care. *JAMA.* Vol 299; 2008:2319-2321.
7. U.S. Department of Health and Human Services. Health Information Technology for the Future of Health and Care. Office of the National Coordinator for Health Information Technology. http://healthit.hhs.gov/portal/server.pt?open=512&mode=2&cached=true&objID=1200. Accessed February 10, 2010.
8. Murphy SN, Mendis ME, Berkowitz DA, Kohane I, Chueh HC. Integration of clinical and genetic data in the i2b2 architecture. *AMIA Annu Symp Proc.* 2006:1040.
9. Murphy SN, Mendis ME, Hackett K, Kuttan R, Pan W, Phillips LC, Gainer V, Berkowicz D, Glaser JP, Kohane I, Chueh HC. Architecture of the open-source clinical research chart from Informatics for Integrating Biology and the Bedside. *AMIA Annu Symp Proc.* 2007;11:548-552.
10. Oster S, Langella S, Hastings S, Ervin D, Madduri R, Phillips J, Kurc T, Siebenlist F, Covitz P, Shanbhag K, Foster I, Saltz J. caGrid 1.0: an enterprise Grid infrastructure for biomedical research. *J Am Med Inform Assoc.* 2008;15(2):138-149.
11. Weber GM, Murphy SN, McMurry AJ, Macfadden D, Nigrin DJ, Churchill S, Kohane IS. The Shared Health Research Information Network (Shrine): a prototype federated query tool for clinical data repositories. *J Am Med Inform Assoc.* 2009;16(5):624-630.
12. Deshmukh VG, Meystre SM, Mitchell JA. Evaluating the informatics for integrating biology and the bedside system for clinical research. *BMC Med Res Methodol.* 2009;9:70.
13. Eccles M, Mittman B. Welcome to Implementation Science. *Implement Sci.* 2006;1(1):1.
14. Colletti RB, Baldassano RN, Milov DE, Margolis PA, Bousvaros A, Crandall WV, Crissinger KD, D'Amico MA, Day AS, Denson LA, Dubinsky M, Ebach DR, Hoffenberg EJ, Kader HA, Keljo DJ, Leibowitz IH, Mamula P, Pfefferkorn MD, Qureshi MA, Improvement. Pediatric IBD Network for research and. Variation in care in pediatric Crohn disease. *J Pediatr Gastroenterol Nutr.* 2009;49(3):297-303.
15. Kappelman M, Crandall W, Colletti R, Leibowitz I, Duffy L, Milov D, Kim S, Schoen B, Patel A, Grunow J, Goudie A, Larry E, Fairbrother G, Margolis P. A short pediatric Crohn's disease activity index for quality improvement and observational research. *Inflamm Bowel Dis.* 2009;15(s2):S1-S65.
16. Olsen L., Aisner D., McGinnis J.M. *The Learning Healthcare System*. Washington, DC: National Academy Press; 2007.
17. Manion FJ, Robbins RJ, Weems WA, Crowley RS. Security and privacy requirements for a multi-institutional cancer research grid: an interview-based study. *BMC Med Inform Decis Mak.* 2009;9:31.
18. Toronto International Data Release Workshop Authors, Birney E, Hudson TJ, Green ED, Gunder C, Eddy S, Rogers J, Harris JR, Ehrlich SD, Apweiler R, Austin CP, Berglund L, Bobrow M, Bountra C, Brookes AJ, Cambon-Thomsen A, Carter NP, et al. Prepublication data sharing. *Nature.* 2009;461(7261):168-170.
19. Nelson B. Data sharing: empty archives. *Nature.* 2009;461(7261):160-163.
20. Schofield PN, Bubela T, Weaver T, Portilla L, Brown SD, Hancock JM, Einhorn D, Tocchini-Valentini G, Hrabe de Angelis M, Rosenthal N, CASIMIR Rome Meeting participants. Post-publication sharing of data and tools. *Nature.* 2009;461(7261):171-173.

21. Science Commons. Towards a Science Commons. http://sciencecommons.org/about/towards/. Accessed January 18, 2010.
22. Walsh D, Passerini K, Varshney U, Fjermestad J. Safeguarding patient privacy in electronic healthcare in the USA: the legal view. *Int J Electron Healthc.* 2008;4(3-4):311-326.
23. Sobel R. The privacy rule that's not. *Hastings Cent Rep.* 2007;37(4):40-50.
24. U.S. Department of Health & Human Services. HHS.Gov/Recovery. American Recovery and Reinvestment Act of 2009. Overview. http://www.hhs.gov/recovery/overview/index.html. Accessed March 17, 2010.
25. Science Commons. http://sciencecommons.org/. Accessed February 5, 2010.
26. Wilbanks J, Boyle J. Introduction to Science Commons. September 7, 2006; http://sciencecommons.org/wp-content/uploads/ScienceCommons_Concept_Paper.pdf. Accessed January 18, 2010.
27. Metzger J, First Consulting Group. *Using computerized registries in chronic disease care.* Oakland, CA: California HealthCare Foundation;February 2004. Available at: http://www.chcf.org/documents/chronicdisease/ComputerizedRegistriesInChronicDisease.pdf.
28. Gliklich RE, Dreyer NA. *Registries for Evaluating Patient Outcomes: A User's Guide. (Prepared by Outcome DEcIDE Center [Outcome Sciences, Inc. dba Outcome] under Contract No. HHSA290200500351 T01).* Rockville, MD: Agency for Healthcare Research and Quality;April 2007.
29. National Institutes of Health, Office of Rare Disease Research. Rare diseases and related terms. http://rarediseases.info.nih.gov/RareDiseaseList.aspx. Accessed January 14, 2010.
30. Brown J, Holmes J, Maro J, Syat B, Lane K, Lazarus R, Platt R, Research from the Developing Evidence to Inform Decisions about Effectiveness (DEcIDE) Network. *Design Specifications for Network Prototype and Cooperative to Conduct Population-Based Studies and Safety Surveillance*.July 2009. Available at: http://effectivehealthcare.ahrq.gov/ehc/products/54/150/2009_0728DEcIDE_DesignSpecNetCoopPopSafety.pdf.
31. caBIG Strategic Planning Workspace. The Cancer Biomedical Informatics Grid (caBIG): infrastructure and applications for a worldwide research community. *Stud Health Technol Inform.* 2007;129(Pt 1):330-334.
32. Loeffler M, Winter A. Editorial: networking clinical research. *Methods Inf Med.* 2007;46(5):572-573.
33. Maro JC, Platt R, Holmes JH, Strom BL, Hennessy S, Lazarus R, Brown JS. Design of a national distributed health data network. *Ann Intern Med.* 2009;151(5):341-344.
34. Pace WD, Cifuentes M, Valuck RJ, Staton EW, Brandt EC, West DR. An electronic practice-based network for observational comparative effectiveness research. *Ann Intern Med.* 2009;151(5):338-340.
35. U.S. Department of Health and Human Services, Office of the National Coordinator for Health Information Technology, CONNECT Community Portal. About CONNECT. http://www.connectopensource.org/about/what-is-CONNECT. Accessed March 26, 2010.
36. Los RK, van Ginneken AM, van der Lei J. Open SDE: a strategy for expressive and flexible structured data entry. *Int J Med Inform.* 2005;74(6):481-490.
37. Venema AC, van Ginneken AM, de Wilde M, Bogers AJ. Is openSDE an alternative for dedicated medical research databases? An example in coronary surgery. *BMC Med Inform Decis Mak.* 2007;7:31.
38. Chapman WW, Cohen KB. Current issues in biomedical text mining and natural language processing. *J Biomed Inform.* 2009;42(5):757-759.
39. Aronsky D, Haug PJ. Assessing the quality of clinical data in a computer-based record for calculating the pneumonia severity index. *J Am Med Inform Assoc.* 2000;7(1):55-65.
40. Brennan PF, Stead WW. Assessing data quality: from concordance, through correctness and completeness, to valid manipulatable representations. *J Am Med Inform Assoc.* 2000;7(1):106-107.
41. Hogan WR, Wagner MM. Accuracy of data in computer-based patient records. *J Am Med Inform Assoc.* 1997;4(5):342-355.
42. Stein HD, Nadkarni P, Erdos J, Miller PL. Exploring the degree of concordance of coded and textual data in answering clinical queries from a clinical data repository. *J Am Med Inform Assoc.* 2000;7(1):42-54.
43. Weiner MG, Embi PJ. Toward reuse of clinical data for research and quality improvement: the end of the beginning? *Ann Intern Med.* 2009;151(5):359-360.
44. NIH Office of Rare Diseases. Rare diseases and related terms.

http://rarediseases.info.nih.gov/RareDiseaseList.aspx. Accessed March 27, 2010.
45. Horbar JD, Rogiwski J, Plsek PE, Delmore P, Edwards WH, Hocker J, Kantak AD, Lewallen P, Lewis W, Lewit E, McCarroll CJ, Mujsce D, Payne NR, Shiono P, Leahy K, Carpenter JH. Collaborative quality improvement for neonatal intensive care. NIC/Q Project Investigators of the Vermont Oxford Network. *Pediatrics.* 2001;107(1):14-22.
46. Simone JV, Lyons J. Superior cancer survival in children compared to adults: a superior system of cancer care? *Background paper, National Cancer Policy Board* 1998; http://www.simoneconsulting.com/PDF/children_survival.pdf. Accessed January 18, 2010.
47. Margolis P, Provost LI, Schoettker PJ, Britto MT. Quality improvement, clinical research and quality improvement research: opportunities for integration. *Pediatr Clin N Am.* 2009;56(4):831-841.
48. Von Hippel E. *Democratizing Innovation.* Cambridge, MA: MIT Press; 2005.
49. Mackner LM, Sisson DP, WV Crandall. Review: psychosocial issues in pediatric inflammatory bowel disease. *J Pediatr Psychol.* 2004;29(4):243-257.
50. Gupta N, SA Cohen, Bostrom AG, Kirschner BS, Baldassano RN, Winter HS, Ferry GD, Smith T, Abramson O, Gold BD, Hayman MB. Risk factors for initial surgery in pediatric patients with Crohn's disease. *Gastroenterology.* 2006;130(4):1069-1077.
51. Markowitz J, Grancher K, Kohn N, Lesser M, Daum F. A multicenter trial of 6-mercaptopurine and prednisone in children with newly diagnosed Crohn's disease. *Gastroenterology.* 2000;119(4):895-902.
52. Turner D, Grossman AB, Rosh J, Kugathasan S, Baldassano R, Griffiths AM. Methotrexate following unsuccessful thiopurine therapy in pediatric Crohn's disease. *Am J Gastroenterol.* 2007;102(12):2804-2812.
53. Hyams J, Crandall W, Kugathasan S, Griffiths A, Olson A, Johanns J, Liu G, Travers S, Heischkel R, Markowitz J, Cohen S, Winter H, Veereman-Wauters G, Ferry G, Baldassano R, REACH Study Group. Induction and maintence of infliximab therapy for the treatment of moderate-to-severe Crohn's disease in children. *Gastroenterology.* 2007;132(3):863-873.
54. Wolters FL, Russel MG, Stockbrugger RW. Systematic review: has disease outcome in Crohn's disease changed during the last four decades? *Aliment Pharmacol Ther.* 2004;20(5):483-496.
55. Kappelman MD, Bousvaros A, Hyams J, Markowitz J, Pfefferkorn M, Kugathasan S, Rosh J, Otley A, Mack D, Griffiths A, Evens J, Grand R, Langton C, Kleinman K, Finkelstein JA. Intercenter variation in initial management of children with Crohn's disease. *Inflamm Bowel Dis.* 2007;13(7):890-895.
56. Lavori PW, Dawson R. Adaptive treatment strategies in chronic disease. *Annu Rev Med.* 2008;59:443-453.
57. Lichtenstein GR, Feagan BG, Cohen RD, Salzberg BA, Diamond RH, Chen DM, Pritchard ML, Sandborn WJ. Serious infections and mortality in association with therapies for Crohn's disease. *Clin Gastroenterol Hepatol.* 2006;4(5):621-630.
58. Marehbian J, Arrighi HM, Hass S, Tian H, Sandborn WJ. Adverse events associated with common therapy regimens for moderate-to-severe Crohn's disease. *Am J Gastroenterol.* 2009;104(10):2524-2533.
59. Siegel CA, Marden SM, Persing SM, Larson RJ, Sands BE. Risk of lymphoma associated with combination anti-tumor necrosis factor and immunomodulator therapy for the treatment of Crohn's disease: a meta-analysis. *Clin Gastroenterol Hepatol.* 2009;7(8):874-881.
60. Punati J, Markowitz J, Lerer T, Hyams J, Kugathasan S, Griffiths A, Otley A, Rosh J, Pfefferkorn M, Mack D, Evans J, Bousvaros A, Moyer MS, Wyllie R, Oliva-Hemker M, Mezoff A, Leleiko N, Keljo D, Crandall W, Pediatric IBD Collaborative Research Group. Effect of early immunomodulator use in moderate to severe pediatric Crohn disease. *Inflamm Bowel Dis.* 2008;14(7):949-954.
61. Van Limbergen J, Russell RK, Drummond HE, Aldhous MC, Round NK, Nimmo ER, Smith L, Gillett PM, McGrogan P, Weaver LT, Bisset WM, Mahdi G, Arnott ID, Satsangi J, Wilson DC. Definition of phenotypic characteristics of childhood-onset inflammatory bowel disease. *Gastroenterology.* 2008;135(4):1114-1122.
62. Vernier-Massouille G, Balde M, Salleron J, Turck D, Dupas JL, Mouterde O, Merle V, Salomez JL, Branche J, Marti R, Lerebours E, Cortot A, Gower-Rousseau C, Colombel JF. Natural history of pediatric Crohn's disease: a population-based cohort study. *Gastroenterology.* 2008;135(4):1106-1113.
63. D'Haens GR, Baert F, van Assche G, Caenepeel P, Vergauwe P, Tuynman H, De Vos M, Can Deventer S, Stitt L, Donner A, Vermeire S, Van de Mierop FJ, Coche JC, van der Woude J, Ochsenkuhn T, van Bodegraven AA, Van Hootegem PP, Lambrecht GL, Mana F, Rutgeerts PJ,

Feagan BG, Hommes D, Belgian Inflammatory Bowel Disease Research Group, North-Holland Gut Club. Early combined immunosuppression or conventional management in patients with newly diagnosed Crohn's disease: an open randomized trial. *Lancet.* 2008;371(9613):660-667.

64. Sandborn WJ. A randomized, double-blind trial comparing infliximab and infliximab plus azathioprine to azathioprine in patients with Crohn's disease naive to immunomodulators and biologic therapy. *Gastroenterology.* 2009;136(5 Suppl 1):A116.

65. Lichtenstein GR, Yan S, Bala M, Blank M, Sands BE. Infliximab maintenance treatment reduces hospitalizations, surgeries, and procedures in fistulizing Crohn's disease. *Gastroenterology.* 2005;128(4):862-869.

66. Clinical Data Interchange Standards Consortium. CDISC: Mission and Principles. http://www.cdisc.org/mission-and-principles. Accessed February 5, 2010.

67. Society for Clinical Data Management. *Good Clinical Data Management Practices*. Available at: http://www.scdm.org/publications/.

68. Moen RD, Nolan TW, Provost LP. *Quality Improvement through Planned Experimentation.* NY: McGraw-Hill; 1999.

69. James B. Information system concepts for quality measurement. *Med Care.* 2003;41(1 Suppl):171-179.

70. Kilo CM. Improving care through collaboration. *Pediatrics.* 1999;103:384-393.

71. Bodenheimer T, Wagner EH, Grumbach K. Improving primary care for patients with chronic illness. *JAMA.* 2002;288(14):1775-1779.

72. Bodenheimer T, Wagner EH, Grumbach K. Improving primary care for patients with chronic illness: the Chronic Care Model, Part 2. *JAMA.* 2002;288(15):1909-1914.

73. Wagner EH. Chronic disease management: what will it take to improve care for chronic illness? *Eff Clin Pract.* 1998;1(1):2-4.

74. Wagner EH, Austin BT, Davis C, Hindmarsh M, Schaefer J, Bonami A. Improving chronic illness care: translating evidence into action. *Health Aff (Millwood).* 2001;20(6):64-78.

75. Wagner EH, Austin BT, Vonkorff M. Organizing care for patients with chronic illness. *Milbank Q.* 1996;74(4):511-554.

76. Hyams JS, Ferry GD, Mandel FS, Gryboski JD, Kibort PM, Kirschner BS, Griffiths AM, Katz AJ, Grand RJ, Boyle JT. Development and validation of a pediatric Crohn's disease activity index. *J Pediatr Gastroenterol Nutr.* 1991;12(4):439-447.

77. Crandall WV, Kappelman M, Colletti RB, et al. Increased remission in a quality improvement collaborative for pediatric Crohn Disease. Paper presented at: NASPHGAN Annual Meeting2009; Washington, DC.

78. Rothman R, DeWalt DA, Malone R, Bryant B, Shintani A, Crigler B, Weinberger M, Pignone M. Influence of patient literacy on the effectiveness of a primary-care based diabetes management program. *JAMA.* 2004;292:1711-1716.

79. Rothman R, Malone R, Bryant B, Horlen C, DeWalt D, Pignone M. The relationship between literacy and glycemic control in a diabetes disease-meangement program. *Diabetes Educ.* 2004;30:263-273.

80. Rothman RL, Malone R, Bryant B, Shintani AK, Crigler B, DeWalt DA, Dittus RS, Weinberger M, Pignone M. A randomized trial of a primary care-based, disease management program to improve cardiovascular risk factors and glycated hemoglobin in patients with diabetes. *Am J Med.* 2005;118:276-284.

81. DeWalt DA, Pignone M, Malone R, Rawls C, Kosnar MC, George G, Bryant B, Rothman RL, Angel B. Development and pilot testing of a disease management program for low literacy patients with heart failure. *Patient Educ Couns.* 2004;55:78-86.

82. DeWalt DA, Malone RM, Bryant ME, Kosnar MC, Corr KE, Rothman RL, Sueta CA, Pignone MP. A heart failure self-management program for patients of all literacy levels: a randomized,controlled trial [ISRCTN11535170]. *BMC Health Serv Res.* 2006;6:30.

83. Berra S, Borrell C, Rajmil L, Estrada MD, Rodriguez M, Riley AW, Forrest CB, Starfield B. Perceived health status and use of healthcare services among children and adolescents. *Eur J Public Health.* 2006;16(4):405-414.

84. Forrest CB, Riley AW, Vivier PM, Gordon NP, Starfield B. Predictors of children's healthcare use: the value of child versus parental perspectives on healthcare needs. *Med Care.* 2004;42(3):232-238.

85. Forrest CB, Starfield B, Riley AW, Kang M. The impact of asthma on the health status of adolescents. *Pediatrics.* 1997;99(2):E1.

86. Forrest CB, Tambor E, Riley AW, Ensminger ME, Starfield B. The health profile of incarcerated male youths. *Pediatrics.* 2000;105(1 Pt 3):286-291.
87. Riley AW, Coghill D, Forrest CB, Lorenzo MJ, Ralston SJ, Spiel G. Validity of the health-related quality of life assessment in the ADORE study: Parent Report Form of the CHIP-Child Edition. *Eur Child Adolesc Psychiatry.* 2006;15 Suppl 1:163-171.
88. Riley AW, Forrest CB, Rebok GW, Starfield B, Green BF, Robertson JA, Friello P. The Child Report Form of the CHIP-Child Edition: reliability and validity. *Med Care.* 2004;42(3):221-231.
89. Riley AW, Forrest CB, Starfield B, Green B, Kang M, Ensminger M. Reliability and validity of the adolescent health profile-types. *Med Care.* 1998;36(8):1237-1248.
90. Riley AW, Forrest CB, Starfield B, Rebok GW, Robertson JA, Green BF. The Parent Report Form of the CHIP-Child Edition: reliability and validity. *Med Care.* 2004;42(3):210-220.
91. Riley AW, Green BF, Forrest CB, Starfield B, Kang M, Ensminger ME. A taxonomy of adolescent health: development of the adolescent health profile-types. *Med Care.* 1998;36(8):1228-1236.
92. Starfield B, Forrest CB, Ryan SA, Riley AW, Ensminger ME, Green BF. Health status of well vs ill adolescents. *Arch Pediatr Adolesc Med.* 1996;150(12):1249-1256.
93. Zerhouni E. Medicine. The NIH Roadmap. *Science.* 2003;302(5642):63-72.
94. Ader DN. Developing the Patient-Reported Outcome Measurement Information System (PROMIS). *Med Care.* 2007;45(5 Suppl 1):S1-S2.
95. Cella D, Yount S, Rothrock N, Gershon R, Cook R, Reeve B, Ader D, Fries JF, Bruce B, Rose M, PROMIS Cooperative Group. The Patient-Reported Outcomes Measurement Information System (PROMIS): progress of an NIH Roadmap cooperative group during its first two years. *Med Care.* 2007;45(5 Suppl 1):S3-S11.
96. Brunelli SM, Joffe MM, Israni RK, Yang W, Fishbane S, Berns JS, Feldman HI. History-adjusted marginal structural analysis of the association between hemoglobin variability and mortality among chronic hemodialysis patients. *Clin J Am Soc Nephrol.* 2008;3(3):777-782.
97. Neugebauer R, van der Laan M, Joffe MM. Causal inference in longitudinal studies with history-restricted marginal models. *Electron J Stat.* 2007;1:119-154.
98. van der Laan MJ, Petersen ML, Joffe MM. History-adjusted marginal structural models and statistically-optimal dynamic treatment regimes. *Int J Biostat.* 2005;1(1):Article 4.
99. Nolan T, Resar R, Haraden C, Griffins FA. *Improving the Reliability of Health Care.* Cambridge, MA: Institute for Healthcare Improvement;2004.
100. Dekker SWA. *Ten Questions about Human Error: A New View of Human Factors and System Safety.* Mahwah, NJ: Lawrence Erlbaum Associates, Inc; 2005.
101. Nunnally JC, Bernstein IH. *Psychometric Theory.* 3rd ed. New York, NY: McGraw-Hill; 1994.
102. Diggle PJ, Heagerty PJ, Liank K, Zeger SL. *Analysis of Longitudinal Data.* 2nd ed. Oxford, IK: Oxford University Press; 2002.
103. Glass GV, Willson VL, Gottman JM. *Design and Analysis of Time-Series Experiments.* Charlotte, NC: Information Age Publishing; 2008.
104. McDowall D, McCleary R, Meidinger E, Hay R. *Interrupted Time Series Analysis.* Beverly Hills, CA: Sage; 1980.
105. Bodenheimer T., Wagner E. H., Grumbach K. Improving primary care for patients with chronic illness: the chronic care model, Part 2. *JAMA.* Oct 16 2002;288(15):1909-1914.
106. Bodenheimer T., Wagner E. H., Grumbach K. Improving primary care for patients with chronic illness. *JAMA.* Oct 9 2002;288(14):1775-1779.
107. Coleman K., Austin B. T., Brach C., Wagner E. H. Evidence on the Chronic Care Model in the new millennium. *Health Aff (Millwood).* Jan-Feb 2009;28(1):75-85.
108. Austin B., Wagner E., Hindmarsh M., Davis C. Elements of Effective Chronic Care: A Model for Optimizing Outcomes for the Chronically Ill. *Epilepsy Behav.* Aug 2000;1(4):S15-S20.
109. Wagner E. H., Austin B. T., Davis C., Hindmarsh M., Schaefer J., Bonomi A. Improving chronic illness care: translating evidence into action. *Health Aff (Millwood).* Nov-Dec 2001;20(6):64-78.
110. Wagner E. H., Austin B. T., Von Korff M. Improving outcomes in chronic illness. *Manag Care Q.* Spring 1996;4(2):12-25.
111. Wagner E. H., Austin B. T., Von Korff M. Organizing care for patients with chronic illness. *Milbank Q.* 1996;74(4):511-544.
112. Kimura J, DaSilva K, Marshall R. Population management, systems-based practice, and planned

chronic illness care: integrating disease management competencies into primary care to improve composite diabetes quality measures. *Dis Manag.* 2008;11(1):13-22.
113. Renders CM, Valk GD, Griffin S, Wagner EH, Eijk JT, Assendelft WJ. Interventions to improve the management of diabetes mellitus in primary care, outpatient and community settings. *Cochrane Database Syst Rev.* 2001(1):CD001481.
114. Holbrook A, Thabane L, Keshavjee K, Dolovich L, Bernstein B, Chan D, Troyan S, Foster G, Gerstein H, for the COMPETE II Investigators. Individualized electronic decision support and reminders to improve diabetes care in the community: COMPETE II randomized trial. *CMAJ.* 2009;181(1-2):37-44.
115. Costello I., Wong I. C., Nunn A. J. A literature review to identify interventions to improve the use of medicines in children. *Child Care Health Dev.* Nov 2004;30(6):647-665.
116. Fisher L., Weihs K. L. Can addressing family relationships improve outcomes in chronic disease? Report of the National Working Group on Family-Based Interventions in Chronic Disease. *J Fam Pract.* Jun 2000;49(6):561-566.
117. Langley GJ, Nolan KM, Nolan TW, Norman CL, Provost LP. The Improvement Guide. A Practical Approach to Enhancing Organizational Performance. San Francisco: Jossey-Bass; 1996.
118. von Hippel E. *Democratizing Innovation*. Cambridge, MA: MIT Press; 2005.
119. Pollart S. M., Elward K. S. Overview of changes to asthma guidelines: diagnosis and screening. *Am Fam Physician.* May 1 2009;79(9):761-767.
120. Sandelowski M. Sample size in qualitative research. *Res Nurs Health.* 1995;18(2):179-183.
121. Ulin P, Robinson E, Tolley E. *Quality Methods in Public Health: A field Guide for Applied Research.* San Francisco, CA: Jossey-Bass; 2005.
122. Strauss A, Corbin J. *Qualitative Research: Grounded Theory Procedures and Techniques.* Newbury Park, CA: Sage Publications, Inc.; 1990.
123. Hibbard JH, Mahoney ER, Stockard J, Tusler M. Development and testing of a short form of the patient activation measure. . *Health Serv Res.* 2005;40(6 Pt 1):1918-1930.
124. Campbell M, Fitzpatrick R, Haines A, Kinmonth AL, Sandercock P, Spiegelhalter D, Tyrer P. Framework for design and evaluation of complex interventions to improve health. *BMJ.* 2000;321:694-696.
125. Lavori PW, Dawson R. Dynamic treatment regimes: practical design. *Clin Trials.* 2004;1(1):9-20.
126. Gordis L. *Epidemiology*. 4th ed. Philadelphia, PA: WB Saunders Co; 2008.
127. Griffiths AM, Otley AR, Hyams J, Quiros AR, Grand RJ, Bousvaros A, Feagan BG, Ferry GR. A review of activity indices and end points for clinical trials in children with Crohn's disease. *Inflamm Bowel Dis.* 2005;11(2):185-196.
128. Walker-Smith JA, Lebenthal E, Branski D. *Pediatric and Inflammatory Bowel Disease: Perspective and Consequences.* Basel; New York: Karger; 2009.
129. Markowitz J, Hyams J, Mack D, Leleiko N, Evans J, Kugathasan S, Pfefferkorn M, Mezoff A, Rosh J, Tolia V, Otley A, Griffiths A, Moyer MS, Oliva-Hemker M, Wyllie R, Rothbaum R, Bousvaros A, Del Rosario JF, Hale S, Lerer T, Pediatric IBD Collaborative Research Group. Corticosteroid therapy in the age of infliximap: acute and 1-year outcomes in newly diagnosed children with Crohn's disease. *Clin Gastroenterol Hepatol.* 2006;4(9):1124-1129.
130. Munkholm P, Langholz E, Davidsen M, Binder V. Frequency of glucocorticoid resistance and dependency in Crohn's disease. *Gut.* 1994;35(3):360-362.
131. Paerregaard A, Uldall Urne F. Anthropometry at the time of diagnosis in Danish children with inflammatory bowel disease. *Acta Paediatr.* 2005;94(11):1682-1683.
132. Wine E, Reif SS, Leshinsky-Silver E, Weiss B, Shaoul RR, Shamir R, Wasserman D, Lerner A, Boaz M, Levine A. Pediatric Crohn's disease and growth retardation: the role of genotype, phenotype, and disease severity. *Pediatrics.* 2004;114(5):1281-1286.
133. Marcus SB, Strople JA, Neighbors K, Weissberg-Benchell J, Nelson SP, Limbers SC, Varni JW, Alonso EM. Fatigue and health-related quality of life in pediatric inflammatory bowel disease. *Clin Gastroenterol Hepatol.* 2009;7(5):554-561.
134. Griffiths AM, Nicholas D, Smith C, Munk M, Stephens D], Durno C, Sherman PM. Development of a quality-of-life index for pediatric inflammatory bowel disease: dealing with differences related to age and IBD type. *J Pediatr Gastroenterol Nutr.* 1999;28(4):S46-S52.
135. Otley A, Smith C, Nicholas D, Munk M, Avolio J, Sherman PM, Griffiths AM. The IMPACT questionnaire: a valid measure of health-related quality of life in pediatric inflammatory bowel disease. *J*

Pediatr Gastroenterol Nutr. 2002;35(4):557-563.
136. Perrin JM, Kuhlthau K, Chughtai ARomm, D, Kirschner BS, Ferry GD, Cohen SA, Gold BD, Heyman MB, Baldassano RN, Winter HS. Measuring quality of life in pediatric patients with inflammatory bowel disease: psychometric and clinical characteristics. *J Pediatr Gastroenterol Nutr.* 2008;46(2):164-171.
137. Feagan BG, Panaccione R, Sandborn WJ, D'Haens GR, Schreiber S, Rutgeerts PJ, Loftus Jr EV, Lomax KG, Yu AP, Wu EQ, Chao J, Mulani P. Effects of adalimumab therapy on incidence of hospitalization and surgery in Crohn's disease: results from the CHARM study. *Gastroenterology.* 2008;135(5):1493-1499.
138. Forrest CB, Lemke KW, Bodycombe DP, Weiner JP. Medication, diagnostic, and cost information as predictors of high-risk patients in need of care management. *Am J Manag Care.* 2009;15(1):41-48.
139. Forrest CB, Weiner JP, Fowles J, Vogeli C, Frick KD, Lemke KW, Starfield B. Self-referral in point-of-service health plans. *JAMA.* 2001;285(17):2223-2231.
140. Iezzoni LI. *Risk Adjustment for Measuring Health Care Outcomes.* 3rd ed. Chicago, IL: Health Administration Press; 2003.
141. Satsangi J, Silverberg MS, Vermeire S, Colombel JF. The Montreal classification of inflammatory bowel disease: controversies, consensus, and implications. *Gut.* 2006;55(6):749-753.
142. Fagan EA, Dyck RF, Maton RN. Serum levels of C-reactive protein in Crohn's disease and ulcerative colitis. *Eur J Clin Invest.* 1982;12(4):351-359.
143. Sands BE, Abreu MT, Ferry GD, Griffiths AM, Hanauer SB, Isaacs KL, Lewis JD, Sandborn WJ, Steinhart AH. Design issues and outcomes in IBD clinical trials. *Inflamm Bowel Dis.* 2005;11 Suppl 1:S22-S28.
144. Hyams J, Markowtiz J, Otley A, Rosh J, Mack D, Bousvaros A, Kugathasan S, Pfefferkorn M, Tolia V, Evans J, Treem W, Wyllie R, Rothbaum R, del Rosario J, Katz A, Mezoff A, Oliva-Hemker M, Lerer T, Griffiths A, Group Pediatric Inflammatory Bowel Disease Collaborative Research. Evaluation of the pediatric crohn disease activity index: a prospective multicenter experience. *J Pediatr Gastroenterol Nutr.* 2005;41(4):416-421.
145. Forrest CB, Nutting PA, von Schrader S, Rohde C, Starfield B. Primary care physician specialty referral decision making: patient, physician, and health care system determinants. *Med Decis Making.* 2006;26(1):76-85.
146. Johnson ML, Crown W, Martin BC, Dormuth CR, Siebert U. Good research practices for comparative effectiveness research: analytic methods to improve causal inference from non-randomized studies of treatment effects using secondary data sources: the ISPOR Good Research Practices for Retrospective Database Analysis Task Force report - Part III. *Value Health.* September 29, 2009;Epub ahead of print.
147. Robins JM. Marginal structural models. Paper presented at: Proceedings of the American Statistical Association. Section on Bayesian Statistical Science.1997; Alexandria, VA.
148. Hernan MA, Hernandez-Diaz S, Robins JM. A structural approach to selection bias. *Epidemiology.* 2004;15(5):615-625.
149. Suarez D, Haro JM, Novick D, Ochoa S. Marginal structural models might overcome confounding when analyzing multiple treatment effects in observational studies. *J Clin Epidemiol.* 2008;61(6):525-530.
150. Robins JM, Hernan MA, Brumback B. Marginal structural models and causal inference in epidemiology. *Epidemiology.* 2000;11(5):550-560.
151. Mortimer KM, Neugebauer R, van der Laan M, Tager IB. An application of model-fitting procedures for marginal strucutral models. *Am J Epidemiol.* 2005;162(4):382-388.
152. Banerjee S, Bishop WP. Evolution of thiopurine use in pediatric inflammatory bowel disease in an academic center. *J Pediatr Gastroenterol Nutr.* 2006;43(3):324-330.
153. Rosenbaum PR, Silber JH. Matching and thick description in an observational study of mortality after surgery. *Biostatistics.* 2001;2(2):217-232.
154. Silber JH, Lorch SA, Rosenbaum PR, Medoff-Cooper B, Bakewell-Sachs S, Millman A, Mi L, Even-Shoshan O, Escobar GT. Time to send the preemie home? Additional maturity at discharge and subsequent health care costs and outcomes. *Health Serv Res.* 2009;44(2 Pt 1):444-463.
155. Silber JH, Rosenbaum PR, Trudeau ME, Even-Shoshan O, Chen W, Zhang X, Mosher RE. Multivariate matching and bias reduction in the surgical outcomes study. *Med Care.* 2001;39(10):1048-1064.
156. Hernan MA, Alonso A, Logan R, Grodstein F, Michels KB, Willett WC, Manson JE, Robbins JM. Observational studies analyzed like randomized experiments: an application to postmenopausal

hormone therapy and coronary heart disease. *Epidemiology.* 2008;19(6):766-779.
157. Brookhart MA, Wang PS, Solomon DH, Schneeweiss S. Evaluating short-term drug effects using a physician-specific prescribing preference as an instrumental variable. *Epidemiology.* 2006;17(3):268-275.
158. Schneeweiss S, Seeger JD, Landon J, Walker AM. Aprotinin during coronary-artery bypass grafting and risk of death. *N Engl J Med.* 2008;358(8):771-783.
159. Leigh JP, Ward MM, Fries JF. Reducing attrition bias with an instrumental variable in a regression model: results from a panel of rheumatoid arthritis patients. . *Stat Med.* 1993;12(11):1005-1018.
160. Rubin DB. *Multiple Imputation for Non-Response in Surveys.* New York, NY: Wiley; 1987.
161. Schafer JL. Multiple imputation: a primer. *Stat Methods Med Res.* 1999;8(1):3-15.
162. Carpenter JR, Kenward MG. *Missing data in randomised controlled trials: a practical guide.* London, UK: Medical Statistics Unit, London School of Hygiene and Tropical Medicine;2007. Available at: http://www.haps.bham.ac.uk/publichealth/methodology/docs/invitations/Final_Report_RM04_JH17_mk.pdf.
163. Berman JJ. A tool for sharing annotated research data: the "Category 0" UMLS (Unified Medical Language System) vocabularies. *BMC Med Inform Decis Mak.* 2003;3:6.
164. Tuttle MS, Nelson SJ. The role of the UMLS in 'storing' and 'sharing' across systems. *Int J Biomed Comput.* 1994;34(1-4):207-237.
165. Benkler Y. *The Wealth of Networks: How Social Production Transforms Markets and Freedom.* New Haven, CT: Yale University Press; 2006.
166. Jones J, Hunter D. Consensus methods for medical and health services research. *BMJ.* 1995;311(7001):376-380.