

ADVANCES IN IBD

Current Developments in the Treatment of Inflammatory Bowel Disease

Section Editor: Stephen B. Hanauer, MD

Using Institutional Databases to Study Inflammatory Bowel Disease



David Binion, MD
 Professor of Medicine
 Co-Director, IBD Center–Translational Research
 University of Pittsburgh
 Pittsburgh, Pennsylvania

G&H What is the value of using an institutional database for research?

DB Historically, it has been difficult for individual health care providers to have a comprehensive quality assessment of their practices. For example, knowing the success rate of treating a certain condition was challenging because providers did not typically keep track of the individual outcomes of patients and then look at them in a collective fashion. Measurement is important in order to achieve quality improvement, which is why the field of gastroenterology is presently being asked to develop quality measures. If providers could track their practices, examine medication rates in their patients, and follow the outcomes of concerning laboratory values, it would be possible to determine when success is achieved, as well as poor outcomes when patients are not doing well. Therefore, the ability to look at data in a comprehensive fashion provides the opportunity to improve the quality of care, optimize patient treatment, and conduct research. In addition, institutional databases provide a platform for discovery, standardization of care, and optimization of health economics. Inflammatory bowel disease (IBD) has embraced evidence-based medicine but still needs more data, particularly in regard to the heterogeneous subgroups of patients who have mild to severe lifetime clinical courses of their illness, to answer many of the questions that are challenging physicians on a routine basis. In my opinion, databases are essential for providing real-world data and helping health care providers answer relevant clinical questions quickly and efficiently.

G&H What are the most significant challenges facing the development of institutional databases?

DB There are numerous challenges facing the development of institutional databases in the United States. First and foremost are regulatory issues. In the spring of 2003, federal regulations for the Health Insurance Portability and Accountability Act (HIPAA) came into effect, so physicians had to specifically obtain patients' permission to look at their medical records for the purpose of research. Thus, HIPAA regulations have to be negotiated before building a database. In addition, an institutional review board (IRB) and research regulatory personnel may feel more comfortable creating a database that is limited in regard to the data being collected. This is because the majority of IRB-approved research studies have a specific hypothesis, and database research initially relies on the creation of the data platform prior to the start of individual projects, which are hypothesis-driven. A very important first step is working with a local IRB to devise an acceptable strategy to have patients consent to be part of a research registry.

Another challenge is ensuring a secure computer environment to house patient data. One solution is to create de-identified registries, in which protected health information (PHI) is removed. This type of data does not need to be housed in a secure computer environment. Generating de-identified data will typically involve computer informatics personnel who work with electronic medical records (EMRs). Physicians can then look at these de-identified data to specifically query snapshots of clinical practice. For the purpose of quality improvement initiatives for improving patient care, de-identified registries are usually acceptable. However, when a research component is involved, the local IRB should be consulted.

Another strategy, which is used by the University of Pittsburgh, involves partnering with information technology experts to create a secure computer share drive, which

is used to house data derived from the EMRs of patients participating in the IBD registry. These registry data are housed on a secure server behind a firewall, where patient-related electronic information is maintained. Database maintenance and research are conducted by individuals who have completed their research compliance modules, and coinvestigators on the IRB protocol are then given network drive access to the secure share drive, where information is stored as part of an IBD registry. When developing and maintaining a database, privacy restrictions are essential and should be resolved on a case-by-case basis.

G&H How common is database use? Do most institutions have their own IBD databases by now?

DB It is unknown how many IBD patient databases are currently being used or how much information is being curated at various centers. Cardiology investigators have made tremendous progress in registry-based research; in comparison, the IBD research community is just beginning to develop this type of natural history investigation. Population-based IBD databases have existed in the United States for some time; the best example is the Olmsted County, Minnesota registry, which has tracked the population of Olmsted County for over 50 years. The Manitoba IBD registry, headed by Dr Charles Bernstein, has been very productive over the past 20 years. There have been many publications based upon the data from these 2 population-based registries. Newer population-based registries include OSCCAR (Ocean State Crohn's and Colitis Area Registry), which was developed by Dr Bruce Sands. In Europe, there have been important patient registries developed in Copenhagen County, Denmark (led by Dr Pia Munkholm), the IBSEN (Inflammatory Bowel Disease in Southeastern Norway) cohort (led by Dr Morten Vatn), and the EPIMAD registry from northern France (led by Dr Jean-Frederic Colombel), as well as institution-specific registries from Leuven, Belgium, among others.

In North America, administrative databases have played an important role in understanding the natural history of IBD. Administrative databases from the federal government include the Nationwide Inpatient Sample, which contains discharge summaries of 1000 hospitals. This database uses International Classification of Diseases–9 and –10 codes to survey the spectrum of health care activity occurring in the inpatient setting. Large health care systems, including Kaiser Permanente, have also been able to perform administrative database research based upon the large number of people encountered by their hospitals, clinics, and providers.

There are other unique, epidemiologic databases that have provided insight into IBD. Dr Ashwin Ananthakrishnan from Massachusetts General Hospital and Harvard

Medical School has utilized the Nurses' Health Study, which includes 2 long-term epidemiologic natural history studies comprising over 100,000 women and a third study that is currently recruiting individuals. There are individuals within these cohorts who have developed IBD during the observational time period. In the first 2 studies, almost 500 patients were diagnosed with IBD, which allowed for exploration of exposure histories, dietary and lifestyle factors, and disease pathogenesis.

Institution-specific IBD databases are unique because there is no standard format, although there has recently been an effort to standardize clinical encounters to create multicenter, real-world observational data. For example, the Crohn's and Colitis Foundation of America (CCFA) is currently making a major initiative to standardize a form that would be applied across a large number of IBD centers for optimization of care and research.

G&H What are other examples of IBD databases, particularly institution-specific ones?

DB There are numerous institution-specific IBD databases in the United States. Examples include those from the University of Chicago, Mount Sinai, Cedars-Sinai Medical Center, the University of Maryland, and the Medical College of Wisconsin. More recently, the Partners HealthCare database, which encompasses Massachusetts General Hospital and Brigham and Women's Hospital, has used a de-identified strategy with the program I2-B2, which was developed by Dr Isaac Kohane. Dr Ananthakrishnan has used this multi-institutional de-identified platform along with a consented, prospective IBD registry called PRISM from Massachusetts General Hospital to explore questions regarding the natural history and clinical trajectory of patients with IBD.

The University of Pittsburgh IBD database, of which I am the co-director, takes a granular approach, whereby physicians have complete access to detailed patient information, including all EMRs, to try to understand subtle and nuanced aspects of the natural history and clinical care of IBD.

G&H What types of information are tracked?

DB Every aspect of information included in an EMR has the potential to become raw data for our prospective registry. We can then analyze the patterns of organized raw data, or metadata, in a comprehensive fashion. This curated metadata can rapidly discern patterns that can identify at-risk populations and provide insight into clinical challenges. At my institution, the database includes biochemical markers of inflammation as well as disease activity and quality-of-life scores, which quantify clinical status in an objective fashion. In addition, we have standardized laboratory parameters and pathology data and time-stamped

records of all prescription data and health care utilization (eg, hospitalizations, surgeries, emergency department visits).

My colleagues and I have also looked at other types of patient data, such as clinic and telephone encounters. We found that telephone calls to our center were a barometer of the patient's clinical status and provided an early warning sign of people who might need to be seen more urgently (ie, people who were heading to the emergency room and possible hospitalization), as these individuals had a markedly increased risk of being hospitalized in the ensuing calendar year.

G&H What are other examples of lessons that have been learned from IBD database research?

DB The difficulty of working with observational data is learning how to handle unstructured clinical information derived from the routine care of patients. Physicians have to learn how to handle multiple observations and different numbers of observations in different patients. Database analysis can simplify and dichotomize data and help inform clinical decision-making in real time. For example, led by Dr Ioannis Koutroubakis, my colleagues and I examined patterns of anemia in IBD patients over a 5-year observational period. Persistent or recurrent anemia turned out to be an objective biomarker of more severe disease, as reflected by higher rates of hospitalization, surgery, and need for escalation of therapy. Because complete blood counts (CBCs) are obtained frequently in IBD patients, we could view multiyear patterns of CBC results in our IBD database. We found that oral iron supplementation was ineffective in the high-risk group and anti-tumor necrosis factor-alpha acceleration therapy did not improve anemia in the majority of patients; intravenous iron replacement therapy plus acceleration of medical therapy was needed.

As the field of gastroenterology moves forward with improved biologic and mechanism-based therapies, it will be essential to define and be able to identify subgroups of patients with IBD, particularly those who have specific cytokines/molecular mechanisms driving their type of IBD inflammation. The best means of doing this would be by keeping track of patients in a comprehensive fashion so that physicians can look at large numbers of individuals over a multiyear time period to obtain a temporal component of the trajectory and natural history of disease. This would allow physicians the opportunity to tease out subgroups that could then be matched for prognostication and optimization of therapy.

G&H How receptive have physicians been with the use of EMRs?

DB I think that many physicians have been resistant to EMRs because this tool requires the recording of information in a regimented fashion and also requires a fair amount

of typing. Historically, the preferred strategy of physicians has involved dictated notes. However, dictated notes do not have as much structure as template EMRs, which can pull together various data sources (eg, vital signs, prescription patterns) into a single document and help standardize information. It can be difficult, particularly when a patient is complex, to systemically go through all of the patient's data in a limited amount of time in an office visit. This is precisely where structured data from EMRs can be very helpful, as the burden of pulling together data is performed by the EMR. This is particularly true for the care of patients with IBD, as physicians cannot use just a single marker to determine how patients are doing; physicians have to search and pull together key pieces of data. EMRs can help physicians do this.

G&H What should be kept in mind when building a database from EMRs?

DB Physicians should start in a modest fashion by first developing lists of patients and their medications. By thinking about groups of patients being treated with various agents, physicians can develop an understanding of these patients. For example, if a person is on immunomodulatory therapy, it might be important to have them undergo laboratory testing on a regular basis. Physicians can easily review the lists to see whether there are any potential problems.

As physicians become more comfortable and adept at using simple databases, they can start tracking more information, such as quality-of-life scores, including individual subscores that quantify fatigue and abdominal pain; disease activity scores; laboratory findings; and prescription exposures. As the field of gastroenterology moves forward, databases will become more complex, and the number of fields and pieces of data tracked will increase.

G&H How representative are databases of the general population?

DB Databases, particularly referral databases, will always be skewed in some regard. Approximately 15% of the IBD population in the United States receives care in high-volume centers, meaning that 85% of IBD patients are treated in community settings, small clinics, and private practices. Referral hospitals, by definition, will always have a more complex patient population. With a database, physicians can see the geographic distribution of their patients and determine whether there are areas that feed into a tertiary referral hospital. Databases are not true population-based cohorts because patients who have mild disease do not necessarily seek care in specialized facilities.

Nevertheless, I think that institution-specific databases can have a fairly representative sample of IBD patients

within a region and can highlight the unmet needs of the disease state. These are the patients who are the most challenging and those who would typically migrate to tertiary or specialized centers. Including these patients can enrich the population that is the focus of clinical and translational investigation, and working with them can generate research data that may ultimately help guide improved IBD care throughout the country.

G&H What is the future of database research in IBD?

DB With databases, physicians have the opportunity to gain insight from the routine care of a patient. The gold standard of evidence-based medicine is the multicenter, randomized, placebo-controlled trial, which provides robust and rich data, but these trials are prohibitively expensive to conduct without dedicated resources, not to mention that it is not possible to conduct an unlimited number of these studies. I think that physicians can gain equal insight into the natural history of disease by looking at populations of patients being treated over time. Right now, physicians are still in the early phases of utilizing databases and EMRs for the purpose of natural history and observational studies of IBD patients. As the field moves forward, it will be easier to pull together information from multiple expert centers or different regions of the country; hopefully, this will allow physicians to provide clinical insight and identify unmet needs where more investigation should be focused. Physicians will have the ability to look at a large number of patients and see any parameters that are unusual or that could be a marker of IBD subtypes that might do poorly.

For example, my colleagues and I have looked at patients with immunoglobulin deficiency, which can mimic Crohn's disease. Patients with immunoglobulin deficiency have a much different natural history and more challenges with infections than patients with Crohn's disease. Thus, it is important to identify these groups as well as their trajectory and natural history and then try to generalize that knowledge to help physicians take better care of patients across the country. This cannot be done in a dedicated trial unless physicians start to understand the heterogeneity of the population being cared for right now.

G&H Are there any ongoing studies involving IBD databases that you are anticipating?

DB A CCFA initiative in the United States is currently following children with IBD from the time of diagnosis in terms of genetic and microbiome characterizations and exposure histories. In Canada, physicians are conducting the GEM (Genes, Environment, and Microbiome) study, which is looking at a population of healthy individuals to

see who ultimately develops IBD to understand the contributing factors of the pathogenesis of the disease.

G&H What are the next steps in research?

DB We need a better understanding of the natural history of IBD and the subgroups of patients currently underneath the umbrella of Crohn's disease and ulcerative colitis. Not all patients respond equally well to the therapies currently available; thus, it would be helpful to develop a means of looking at routine, easily available parameters and then be able to predict response to particular therapies.

In addition, there is a lack of familiarity in the general gastroenterology community about databases, simply because they are still fairly new in terms of routine usage. I think that the administrative datasets that have been explored over the past several years have been helpful at identifying new hypotheses and confirming observations that have been seen clinically. One example is the expansion of *Clostridium difficile* infection in North America over the past 2 decades, which hit the IBD patient population quite hard. Based upon information tracked in institutional databases at Barnes Hospital/Washington University in St Louis and the Medical College of Wisconsin, it became clear to physicians that the usual treatment approach to *C difficile* infection had to change due to higher-than-expected rates of patients becoming ill, being hospitalized, and undergoing colectomy. Physicians were able to confirm what was being seen locally by using the Nationwide Inpatient Sample—that there was a rise of *C difficile* infection across the country, particularly in the IBD population. Thus, databases can provide an important real-world message about the challenges facing patients. Evidence-based medicine from clinical trials is very important, but identifying who will do best with a specific agent can be addressed more effectively using database analysis.

Dr Binion has no relevant conflicts of interest to disclose.

Suggested Reading

Ananthkrishnan AN, Cagan A, Cai T, et al. Identification of nonresponse to treatment using narrative data in an electronic health record inflammatory bowel disease cohort. *Inflamm Bowel Dis*. 2016;22(1):151-158.

Atreja A, Achkar JP, Jain AK, Harris CM, Lashner BA. Using technology to promote gastrointestinal outcomes research: a case for electronic health records. *Am J Gastroenterol*. 2008;103(9):2171-2178.

Issa M, Vijayapal A, Graham MB, et al. Impact of *Clostridium difficile* on inflammatory bowel disease. *Clin Gastroenterol Hepatol*. 2007;5(3):345-351.

Koutroubakis IE, Ramos-Rivers C, Regueiro M, et al. Five-year period prevalence and characteristics of anemia in a large US inflammatory bowel disease cohort [published online October 17, 2015]. *J Clin Gastroenterol*. doi:10.1097/MCG.0000000000000417.

Ramos-Rivers C, Regueiro M, Vargas EJ, et al. Association between telephone activity and features of patients with inflammatory bowel disease. *Clin Gastroenterol Hepatol*. 2014;12(6):986-994.e1.