Published in final edited form as:

Semin Thromb Hemost. 2013 February ; 39(1): 10–14. doi:10.1055/s-0032-1329551.

The Institution-Based Prospective Inception Cohort Study: Design, Implementation, and Quality Assurance in Pediatric Thrombosis and Stroke Research

Timothy J. Bernard, MD^{1,2,3}, Jennifer Armstrong-Wells, MD, MPH^{1,2,3}, and Neil A. Goldenberg, MD, PhD^{4,5}

¹Department of Pediatrics, Section of Child Neurology, University of Colorado and Children's Hospital Colorado, Aurora, Colorado ²The Mountain States Regional Hemophilia and Thrombosis Center, University of Colorado and Children's Hospital Colorado, Aurora, Colorado ³Pediatric Stroke Program, University of Colorado and Children's Hospital Colorado, Aurora, Colorado ⁴Pediatric Thrombosis Program and All Children's Research Institute, All Children's Hospital-Johns Hopkins Medicine, St. Petersburg, Florida ⁵Pediatric Thrombosis Program and Divisions of Hematology, Departments of Pediatrics and Medicine, Johns Hopkins School of Medicine, Baltimore, Maryland

Abstract

The development of well-designed cohort studies in rare diseases can lead to the discovery of new risk factors and prognostic markers, enhance understanding of natural history and outcomes, and provide preliminary data for randomized controlled trials of treatment strategies. Designing a robust cohort requires substantial upfront design and planning. Ideally, a cohort study of diseased individuals follows patients prospectively from the time of diagnosis (i.e., from the disease's inception). The objective of this article is to discuss the design and implementation of an institution-based prospective inception cohort study, with applied examples in pediatric stroke and thrombosis. Furthermore, we will discuss the ongoing management and quality assurance mechanisms necessary to optimize such a study. Although the resources necessary to implement a prospective inception cohort study are large, this approach can provide critical observational evidence on natural history and prognostic factors. Following multicenter validation, its findings can inform the design and execution of much-needed randomized controlled clinical trials.

Keywords

cohort; children; quality assurance; stroke; thrombosis

Copyright © 2013 by Thieme Medical Publishers, Inc.

Address for correspondence and reprint requests Neil A. Goldenberg, MD, PhD, All Children's Hospital-Johns Hopkins Medicine, 501 Sixth Avenue South, St. Petersburg, FL 33701 (Neil.Goldenberg@allkids.org).

Conflicts of Interest

The authors declare that they have no relevant conflicts of interest to disclose.

A cohort study is designed to study a group of individuals with a common disease or a common risk factor within a defined time period. Subjects are followed longitudinally after their exposure (defined as disease onset or risk factor identification), and monitored for outcomes of interest (such as death, disability, or new disease onset). Hence, a prospective cohort study can consist of a “healthy” cohort, an “at risk” cohort (sometimes with challenging distinction between the two), or a “diseased” cohort. As examples from childhood arterial ischemic stroke (AIS), investigators can follow patients from onset of stroke as well as patients without stroke but with risk factors. In childhood AIS, most cohorts are defined by onset of disease—childhood AIS. Enrollment in the cohort begins at stroke onset, and investigators follow patients longitudinally for outcomes of interest such as death, disability, and stroke recurrence.

Traditionally, population-based cohorts are employed using specific geographic regions, such as cities, counties, or states/provinces (e.g., Framingham cohort),¹ and sometimes professions (e.g., Nurses’ Health Study).² In contrast, institution-based cohorts typically enroll from a hospital or health-care system. Both types of studies follow patients over time for outcomes of interest, but each has its own strengths and weaknesses. Population-based cohorts are best suited for identifying risk factors for incident disease, whereas institution-based cohorts are well positioned for determining natural history and outcome predictors among those with established disease. These objectives require capture of treatment variables, markers of disease activity, and other modifying factors not usually available by the large epidemiologic design of a population-based cohort. When so-called hard and reliable endpoints, such as mortality, are of key interest, the larger sample size of the population-based cohort study affords greater power and better generalizability than its institution-based counterpart. A caution, however, is the degree to which capture of so-called soft (e.g., functional or subjective) outcomes and definition of disease are reliable in the population-based cohort.

In rare disease, institution-based retrospective cohort studies and case series are common. Prospective cohort studies, however, collect data forward in time from the point of enrollment, adding cost and time-intensiveness to the study planning and execution. Data quality is superior in the prospective design, in which data fields are uniformly defined and collected, rather than relying upon available records and nonstandardized disease and outcome definitions (see also *quality assurance*, in the following text). When data collection occurs during a subject-encounter, recall bias and missing data are minimized. In a noninception cohort study, however, the time after exposure at which patients begin their prospective follow-up on study is variable, limiting the ability to investigate the relationship between early/acute postexposure factors and longer-term outcomes. For this reason, an inception cohort provides an additional enhancement in data quality. As shown schematically in Fig. 1, in an inception-type prospective cohort, subjects are enrolled within a restricted time period postexposure (e.g., within 2 weeks of hospitalization for presenting illness), and followed forward in time for long-term outcomes. Follow-up can be arranged at specific time points, with specifically attained measures at those time points, minimizing the variability of outcome assessment. Given the high quality of the institution-based prospective inception cohort study, its findings are worthy of direct application in an

expanded multicenter cohort for substantiation of findings, culminating in the design of randomized controlled clinical trials (RCTs) employing prognostic stratification.

Study Design and Setup

In the planning of an institution-based prospective inception cohort study, substantial effort is needed in study design. Clinical and scientific input from a broad multidisciplinary group of collaborators should ideally be sought when developing eligibility criteria, disease definitions, variables of interest and their variable definitions, data collection methods, statistical analysis plans, and logistical plans for study implementation. In the rare disease area of childhood AIS, for example, the collaborative team might consist of neurologists, hematologists, neuroradiologists, neurosurgeons, rehabilitation physicians (i.e., physiatrists), neuropsychologists, cardiologists, and rheumatologists, given the frequent clinical involvement of these specialists in the care of childhood AIS patients. Uniform nomenclature and definitions (e.g., a data glossary) for use in data collection and coding are essential to data quality assurance, particularly amid the anticipated turnover in clinical research personnel in a longitudinal study. Additionally, tools for standardized outcome assessment in the settings of best clinical care as well as research are critical to the reliability of endpoints, and personnel must be adequately trained (and training periodically refreshed) in the use of these instruments. For example, in pediatric venous thrombosis, wherein chronic venous insufficiency (i.e., the postthrombotic syndrome) is an important long-term outcome, standardization of a pediatric postthrombotic syndrome outcome measure has been an important advancement in clinical care and research,³ and a training video has been made available via the Internet.⁴

Partnering with a biostatistician or other clinical research methodologist in the development of testable hypotheses and overall study design will help prevent design pitfalls that otherwise may become apparent only years later. Optimal prospective inception cohort study design, much like that of a clinical trial, includes consideration of power and sample size for each hypothesis, and outlines statistical analysis plans addressing each. Common statistical methods to evaluate exposure-outcome relationships (e.g., putative prognostic factors) include univariate and multivariate logistic regression and/or Cox proportional hazard modeling.

Equally important to the scientific and statistical team is the inclusion of research operations personnel that include a clinical research coordinator and specimen processor. The research operations team is helpful in collaborating with the principal investigator in developing a study Manual of Procedures (MOP; alternatively called a Study Operations Manual or Manual of Operations), to describe all research procedures in detail (see Table 1). The MOP may include procedures for outcome assessments, particularly those for functional endpoints or that involve patient-reported outcomes. In addition, the MOP should describe uniform specimen collection, processing, and storage requirements. The MOP should also address those “routine” clinical procedures that may lack uniformity. For example, in institution-based prospective inception cohort study of pediatric venous thrombosis, the description of ultrasound image acquisition in the MOP is important to quality assurance, much as would be done in a clinical trial; specification of the capture of a video clip of the compression

maneuver in a given venous segment greatly assists in the definitive determination of the presence of thrombosis and can be further ensured through the development of institutional procedures for vascular imaging collaboratively with a radiologist coinvestigator.

The MOP can also provide a helpful guide for data extraction from source documentation in the medical record, as well as provide a glossary for definitions of data fields collected on Case Report Forms (CRFs). For example, in an institution-based prospective inception cohort study of pediatric venous thromboembolism (VTE), how shall *catheter-associated* thrombosis be defined? How will *recurrent VTE* be captured distinctly from *local thrombus progression*, with regard to anatomic presentation and time of onset/diagnosis after qualifying VTE event? What distinguishes *clinically relevant* bleeding episodes from all other bleeds? As is the case in clinical trials, endpoint definitions in the institution-based prospective inception cohort study must be precise, uniformly applied, and clearly described in the Methods section of all publications stemming therefrom. Study findings can best be considered in the context of these definitions, which can then inform observational study selection for inclusion in systematic reviews and meta-analyses in particular. In addition, presenting and interpreting study findings in the context of clear endpoint definitions allows observational data to more reliably inform the design of clinical trials.

Database design is another essential element to successful study execution because spreadsheets have limited functionality and poor security measures. A database designer/manager can assist in development and implementation of a user-friendly, reliable, and compliant method for data storage, with programmed reports. Database design and implementation should recognize Institutional Review Board/Ethics Board and Health Insurance Portability and Accountability requirements for patient confidentiality by ensuring a one-to-one correlation between data fields collected and data fields specified in the protocol/Institutional Review Board/Ethics Board application, including any protected health information. Some databases have the capability of adhering to U.S. Code Federal Regulations part 11 guidelines by maintaining an audit trail; when implemented within appropriate standard operating procedures, part 11 compliance can be ensured.

Ensuring accurate data entry is also an essential part of the study start-up. Research personnel should develop and follow a standardized data quality assurance plan within the study MOP, which typically includes systematic verification against source documents. Database design should also accommodate tracking of quality assurance procedures, as discussed in the following section.

Study Execution

The development of a multidisciplinary clinical team dedicated both to clinical care and the successful conduct of the prospective inception cohort study is invaluable. An invested team of clinicians can ensure patient identification, standardized approach to care, longitudinal follow-up, and proper data acquisition. This team, in partnership with research personnel, should develop a robust and standardized screening system for patient identification, fulfilling both clinical needs and research objectives. As an example from pediatric stroke, the authors and colleagues at Children's Hospital Colorado developed a stroke alert system

that identifies patients with suspected childhood-onset AIS immediately at presentation, using a dedicated stroke team, a standardized care pathway for diagnostic evaluation and acute management, and “best practice alerts” within an electronic health record. This process ensures identification of most patients, minimizing referral bias within the institution and facilitating careful enumeration of screened, eligible, and enrolled populations (so-called denominator issues) for the purposes of cohort study tracking and reporting.

Clinical care pathways for standardized management of the target patient population not only optimize care (e.g., serving as the foundation for quality improvement measures) but permit research questions to be posed around known, measurable variations in care. For example, in childhood-onset AIS, the bleeding risk associated with anticoagulant use in nonmoyamoya arteriopathies was explored at Children’s Hospital Colorado in a mixed retrospective-prospective cohort study in collaboration with colleagues at the University of Münster, in which the research aims relied importantly upon standardized procedures for administering anticoagulation as part of institutional diagnostic classification and management pathways.⁵ This collaborative study provided important preliminary data upon which future randomized trials can be designed. Standardization in radiologic and laboratory testing, both diagnostically and for treatment monitoring, can also be facilitated by institutional care pathways. Although motivated by the need to optimize clinical care, such standardization can also greatly improve the degree of missing data of key covariates in cohort studies. For example, the authors and colleagues recently undertook to evaluate markers of systemic inflammation in childhood-onset AIS,⁶ using clinically obtained C-reactive protein and erythrocyte sedimentation rate measurements as part of routine diagnostic workup for vasculitis in young patients with stroke. During this analysis, it was clear that the degree of missing data on these inflammatory markers from clinical workup substantially declined in the cohort following implementation of an electronic medical record–based order set paired with the institutional clinical care pathway for management of childhood-onset AIS.

Another important element of prospective inception cohort study execution is data quality assurance. Similar to data monitoring visits for industry-sponsored studies, internal audits of CRFs are advised to ensure CRF completeness as well as consistency with source documents. When resources prohibit reviewing all CRFs, monitoring of randomly sampled CRFs enables identification of systematic induced errors. Although costly, dual data entry can provide systematic quality assurance for data integrity beyond internal auditing. Dual data entry should be performed by independent research personnel and allows research staff to easily identify errors by automatically flagging discrepant duplicate entries. When resources are limited, dual data entry can be reserved for key variables of interest to ensure integrity of the most important variables.

Tracking and documentation of adherence to data quality assurance procedures is also important for key data elements of interest or entire data per subject. For example, a data field capturing a check-off procedure (“yes/no”) for data quality assurance can be implemented and time-stamped with expectations for same outlined within the study MOP. In the context of a broad multidisciplinary team, individual team members can be assigned

verification of data relevant to their expertise. For example, in a pediatric VTE cohort, pulmonary embolism subcohort, echocardiographic data extracted from the medical record by the study coordinator/research nurse during a combined clinical and research follow-up visit can be verified (and documented in the database as such) by the cardiology coinvestigator within a preestablished time frame (e.g., 7 to 10 days) from the visit date. For enhanced data quality assurance, routine programmed reports can be used to report “missing/overdue” data elements, including the data verification field, to facilitate follow-up communication between the principal investigator and the cardiology coinvestigator. The same process can be followed for data verification/quality assurance in the following additional examples: radiologic findings on thrombus burden in a limb deep venous thrombosis subcohort (radiology coinvestigator); in a childhood-onset AIS subcohort, radiologic findings for both incident and recurrent childhood-onset stroke, to facilitate standardized classification using the Childhood Arterial Stroke Classification and Diagnostic Evaluation criteria (neuroradiology coinvestigator)⁷; neurologic findings in a cerebral sinovenous thrombosis subcohort (neurology coinvestigator); and rheumatologic laboratory testing data in an antiphospholipid antibody syndrome subcohort (rheumatology coinvestigator). These and other data quality assurance measures are summarized in Table 2.

Conclusions

The development of an institution-based prospective inception cohort study requires significant resources and forethought. The cooperation of team members within and between devoted groups of clinicians and researchers enhances study design and facilitates its seamless execution. An effective model for prospective inception cohort studies borrows operational and quality assurance concepts from the clinical trial, including specific aims and hypotheses, analytic plans, a manual of procedures, uniform follow-up assessments, training in standardized outcome measurement, a data glossary, and data verification procedures. The return on investment of the successful institution-based prospective inception cohort study includes a portfolio of multidisciplinary publications on natural history, prognostic markers, and prognostic (i.e., “risk”) stratification. More importantly, when followed by multicenter prospective cohort studies designed to validate single-institutional findings, the value of the institution-based prospective study is in the establishment of highest-quality observational evidence to inform the development of interventional studies that will improve future patient outcomes.

Acknowledgments

Dr. Bernard is funded in part by a Career Development Award from the National Institutes of Health, National Heart Lung and Blood Institute (1K23HL096895–01). Dr. Armstrong-Wells is funded in part by a Career Development Award from the National Institutes of Health, Building Interdisciplinary Research Careers in Women’s Health (KD HDO57022). Dr. Goldenberg was supported in the development of this manuscript by an Eberhard F. Mammen Excellence in Thrombosis and Haemostasis Award from *Seminars in Thrombosis & Hemostasis*. Concepts discussed in this article were presented by Dr. Goldenberg at the American Society of Hematology Clinical Research Training Institute (La Jolla, CA, United States; August 2010) and the XXIII Congress of the International Society on Thrombosis and Haemostasis (Kyoto, Japan; July 2011).

References

1. Dawber TR, Meadors GF, Moore FE Jr. Epidemiological approaches to heart disease: the Framingham Study. *Am J Public Health Nations Health*. 1951; 41(3):279–281. [PubMed: 14819398]
2. Belanger CF, Hennekens CH, Rosner B, Speizer FE. The nurses' health study. *Am J Nurs*. 1978; 78(6):1039–1040. [PubMed: 248266]
3. Goldenberg NA, Donadini MP, Kahn SR, et al. Post-thrombotic syndrome in children: a systematic review of frequency of occurrence, validity of outcome measures, and prognostic factors. *Haematologica*. 2010; 95(11):1952–1959. [PubMed: 20595095]
4. University of Colorado Denver School of Medicine. [Accessed August 03, 2012] Training video: clinical PTS exam example. <http://www.ucdenver.edu/academics/colleges/medicalschoo/centers/HemophiliaThrombosis/research/kidsdott/Pages/Sponsors%20and%20Links.aspx>
5. Bernard TJ, Goldenberg NA, Tripputi M, Manco-Johnson MJ, Niederstadt T, Nowak-Göttl U. Anticoagulation in childhood-onset arterial ischemic stroke with non-moyamoya arteriopathy: findings from the Colorado and German (COAG) collaboration. *Stroke*. 2009; 40(8):2869–2871. [PubMed: 19478216]
6. Bernard TJ, Fenton LZ, Apkon SD, et al. Biomarkers of hypercoagulability and inflammation in childhood-onset arterial ischemic stroke. *J Pediatr*. 2010; 156(4):651–656. [PubMed: 20022340]
7. Bernard TJ, Manco-Johnson MJ, Lo W, et al. Towards a consensus-based classification of childhood arterial ischemic stroke. *Stroke*. 2012; 43(2):371–377. [PubMed: 22156694]

Fig. 1.

Schema for the prospective inception cohort study design. (Adapted from: Szklo M, Nieto FJ. *Epidemiology: Beyond the Basics*. 2nd edition. Sudbury, MA: Jones and Bartlett; 2007:23).

Table 1

Example “Table of Contents” for a manual of procedures in a prospective inception cohort study of pediatric venous thromboembolism

Section and topic	Page
1 Contact information	1
2 Study documentation	2
2.1 Regulatory binder contents	
2.2 Data collection	3
2.2.1 Overview	3
2.2.2 Schedule of assessments	4
2.2.3 Case report forms	5
2.2.4 Data completion instructions and glossary	15
2.2.4.1 Recurrent VTE definition	19
2.2.4.2 Postthrombotic syndrome definitions	19
2.2.4.3 Bleeding definitions	20
3 Source documentation	21
3.1 Medical history	21
3.2 Physical examination	22
3.3 Diagnostic evaluation	23
3.4 Treatment/medications	24
3.5 Outcome assessments	25
3.6 Other	27
4 Study procedures	28
4.1 Screening	28
4.2 Informed consent process	28
4.3 Enrollment	29
4.4 Withdrawal	29
4.5 Research specimen collection, processing, tracking and storage	30
5 Quality assurance measures	37
5.1 Diagnostic evaluation	37
5.1.1 Imaging procedures: Compression ultrasound with Doppler	38
5.2 Standardized outcome assessment	40
5.2.1 Postthrombotic syndrome	40
6 Adverse events	42
6.1 Definitions	42
6.2 Surveillance	44
6.3 Reporting	45

Abbreviation: VTE, venous thromboembolism.

Table 2

Quality assurance measures in the prospective inception cohort study

Data quality assurance
<ul style="list-style-type: none"> • Involvement of clinical content and clinical science experts in protocol and database design • Diagnostic/eligibility adjudication committee • Data glossary (standardized definitions) • Dual data entry (key fields and endpoints) • Timely data verification by clinical content experts
Endpoint quality assurance
Endpoint selection
<ul style="list-style-type: none"> • Clinically relevant • Standardized • Validated (if available)
Endpoint measurement and reporting
<ul style="list-style-type: none"> • Training (initial and refreshers) on outcome measures/instruments for “soft” endpoints • Endpoint adjudication committee