

Interval and Clinical Cohort Studies: Epidemiological Issues

BRYAN LAU,^{1,2} STEPHEN J. GANGE,² and RICHARD D. MOORE^{1,2}

ABSTRACT

Cohort studies based upon clinic populations and medical records are becoming more abundant due in part to an increasing trend toward electronic medical records and advancement in information technology. This design has been utilized in the HIV setting to great success and involves following individuals as they access medical care. These clinical cohort designs have not been compared to the classic interval cohort design in which individuals are followed at specified intervals that are unrelated to the participants' ongoing health care. The interval and clinical cohort designs are distinguished and the advantages and disadvantages inherent in each design are discussed.

INTRODUCTION

THE COHORT STUDY DESIGN can be defined as following groups of individuals with and without an exposure to determine whether or not an outcome occurs differentially between exposure groups.¹ As reviewed by Liddell² and Samet and Munoz,³ the cohort study design has had an interesting history and has been an invaluable tool for studying numerous diseases.

The current convention is to classify cohort studies according to several attributes of the data and of the population being studied:

Characterization of the timing of follow-up as either historical (also termed retrospective or nonconcurrent) or prospective (or concurrent) designs. The retrospective cohort study is one in which individuals are identified as having an exposure at some point in the past and are followed forward in time to the development of the outcome of interest. An attribute of retrospective cohort studies is that these studies are constructed from data that have previously been collected. Alternatively, prospective designs collect current exposure status and the individuals are then followed into the future. The retrospective and prospective designs do not have to be mutually exclusive.⁴

Characterization of the type of data collected as either life-table or longitudinal designs. Life-table and longitudinal cohorts are broad classifications to characterize how time and exposure are treated.⁵ Life-table cohort studies are those in which exposures at a single time point are collected and linked to subsequent event times.⁵ Cohort studies conducted prior to the

1970s relied heavily on life-table methods for analysis.³ Subsequently, more complicated longitudinal designs that include repeated measurement of exposures have become widely used. These designs reflect the more complex questions of investigating disease etiology and capitalize on the evolution of statistical methods for analyzing repeated observations. Studies with this design can examine changes in exposure status over time and their association with population and individual-level changes of disease over time.⁵

Characterization of study enrollment characteristics as either population based, multicenter, or single center. Population-based studies utilize formal probability sampling methods to select individuals from a population, usually defined by geographic area, during a fixed interval of time.⁶ Sophisticated sampling schemes can be useful to obtain important representation, thereby enhancing calculation of population level summaries. Multicenter and single center cohort studies entail recruiting study subjects from either multiple or single sites to direct recruiting efforts toward select groups of individuals that may not be efficiently observed through population sampling.

Characterization of participant enrollment as either fixed or dynamic studies. Fixed cohorts are studies in which enrollment occurs over a specified time and once completed no additional participants are enrolled. Conversely, dynamic cohorts do not have an enrollment period, allowing individuals to enroll throughout the course of the study. Therefore while the population size of a fixed cohort may only decline over time due to withdrawal from the study (either by loss to follow-up, refusal, or death), the size of a dynamic cohort may fluctuate over time.

¹Division of General Internal Medicine, Johns Hopkins University School of Medicine, and the ²Department of Epidemiology, Johns Hopkins University Bloomberg School of Public Health, Baltimore, Maryland, 21218.

While these traditional criteria are not exhaustive and have been useful for classifying aspects of the cohort design, they fail to consider emerging evolution of cohort studies that is being fueled by advances in information technology. In particular, in this paper we define the clinical cohort study design, distinguish it from interval cohort studies, and outline the epidemiological similarities and differences. These epidemiological issues may highlight potential biases that may affect the interpretation of analyses conducted within cohort studies. This report is the first we are aware of that illustrates the epidemiological issues that may impact the validity of interval and clinical cohort studies. We apply familiar epidemiological concepts to highlight the advantages and disadvantages of these cohort designs.

COHORT CLASSIFICATION

In this paper, cohort studies are classified into two groups: the interval and the clinical cohort. In an interval cohort study, individuals are followed at specified intervals, with data collection under the control of the investigator, independent of the source population.

Several landmark interval cohort studies were initiated in the 1940s and 1950s. The Framingham Study was initiated in the late 1940s to examine cardiovascular disease and has contributed greatly to our understanding of this disease.^{2,3,7} Another landmark study from the 1940s is the Japanese Life Span Study, which has followed the atomic bomb survivors in Hiroshima and Nagasaki. This study has been a primary source for elucidating the affects of acute radiation exposure.^{2,3,8} During the 1950s, the British Physician study was initiated and was instrumental in elucidating the association between smoking and lung cancer.^{2,3,9}

In HIV research there are quite a few interval cohort studies such as the AIDS Link to the Intravenous Experience study,¹⁰ the Amsterdam Cohort Studies,¹¹ the Italian Seroconversion Study,¹² the Multicenter AIDS Cohort Study (MACS),¹³ and the Women's Interagency HIV Study (WIHS).¹⁴ For example, the MACS was initiated in 1983 to study the natural history of HIV infection among homosexual and bisexual men in the United States. Between 1984 and 1985, 4954 HIV⁺ and HIV⁻ men were enrolled in Baltimore/Washington DC, Chicago, Los Angeles, and Pittsburgh. At semiannual visits, participants returned to the clinics to provide specimens for laboratory analyses, undergo a physical examination, and complete self-administered data forms and an interviewer-administered questionnaire.¹³ The WIHS study began 10 years later utilizing a similar design among HIV-infected and at-risk women.¹⁴

The clinical cohort design is also used in HIV research. In contrast to the interval cohort design, this design involves taking either a portion or all of the individuals who are actively receiving health care for some condition² and use the data that are obtained from ongoing care for the patient to assess outcomes. Thus, the source population is a clinic setting, and the type and timing of data collected are entirely determined by the nature of their health care services.

There are several superb examples of clinical cohort studies in HIV research, such as the Johns Hopkins HIV Clinical Practice Cohort (JHH CPC)¹⁵ and the Swiss Cohort.¹⁶ The JHH CPC

is a longitudinal, single site, dynamic, clinical cohort of patients receiving care through the Johns Hopkins AIDS Service, which provides care for a large proportion of HIV-infected patients in the Baltimore metropolitan area.¹⁵ The JHH CPC has followed patients since 1989 with the goal to better understand the effectiveness of therapy in the clinical practice. To date, almost 6000 patients have been enrolled into the cohort. Enrollment into the cohort coincides with first enrollment into the HIV clinic. Approximately 85% of patients receive all of their health care within the Johns Hopkins Health System. Extensive patient information (laboratory, diagnostic, clinical, pharmaceutical, behavioral, and social) is collected at the time of enrollment and entered into a database. Information on patients is collected over time through medical records, the Johns Hopkins Health System automated databases, supplemental medical records from other facilities, vital records, and automated computer-assisted self-interviews. However, unlike the interval cohort design, participants are not brought into the clinic at scheduled visits for the purposes of the study. Rather the information on participants becomes available as they seek medical care.¹⁵ The JHH CPC may obtain an individual's CD4⁺ counts more or less often than the MACS (semiannual) depending on the individual's propensity to regularly obtain medical care and the provider's practice patterns which contributes to routine clinical scheduling (but not necessarily to whether the patient is adherent to scheduled appointments) and dictates whether laboratory testing is conducted.

Clinical cohort studies may also include studies that are conducted in health administration databases. While these data sources were primarily formed for administrative purposes and not for epidemiological studies, they have become a convenient tool for epidemiological research. Several examples of private insurance databases are the Kaiser Permanente Medical Care Program, Harvard Pilgrim Health Care, and United Health Group.¹⁷ Several examples of government-sponsored databases are Medicare, Medicaid, and the United States Renal Data System databases.¹⁷ While the databases are not cohort studies, clinical cohort studies can be formed when investigators assemble and follow individuals based upon exposure information until the event of interest occurs. We classify studies conducted on data from administrative databases as clinical cohorts for two reasons. First, data included in the database are generated during the care of individuals. Second, individuals do not have a structured study interval visit as in the interval cohort design and data are collected based upon their health care utilization. An example of this type of administrative database being used for a clinical HIV cohort is within Kaiser Permanente Medical Care Program of Northern California (KPNC). KPNC has maintained a database of HIV-positive individuals since 1988 in which members obtain most of their medical care, ancillary services, and prescription medications from KPNC facilities. Electronic data are available, which include inpatient and outpatient services and diagnoses along with pharmacy utilization.¹⁸

SIMILARITIES BETWEEN DESIGNS

There are several attributes that are shared by both the interval and clinical cohort design (Fig. 1, Table 1), which include general advantages and disadvantages of cohort studies.

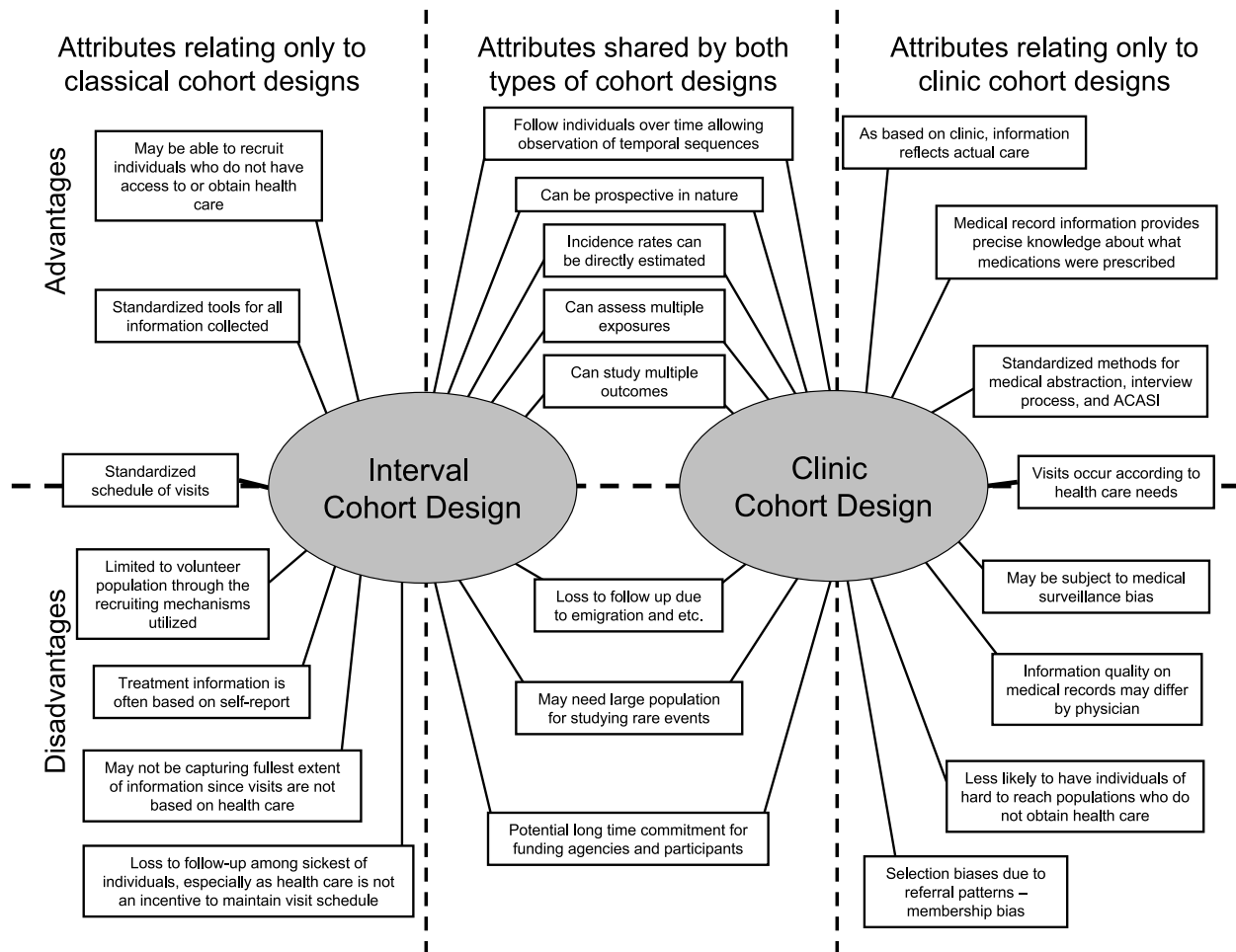


FIG. 1. A schematic of the advantages and disadvantages of the interval and clinical cohort study designs categorized by whether the attributes are shared or exclusive to the cohort design.

All cohort studies are designed to follow individuals over time allowing for observation of temporal sequences of exposures and events, which is a major advantage of cohort studies. An exposure could not be a causal agent for a disease if it occurred

after or as a consequence of the disease. A further advantage of these cohort designs is their prospective nature. The temporal nature of cohort studies minimizes recall bias as the exposure is assessed concurrently with the follow-up of individuals

TABLE 1. ADVANTAGES AND DISADVANTAGES OF COHORT STUDIES

Advantages	
Temporality	Temporal nature of cohort studies allows observation of temporal sequence
May be prospective	Prospective cohort studies (1) do not have to rely on data that have been previously collected and (2) minimize recall bias from participants as exposure and outcomes are being continuously assessed
Incidence	Cohort studies allow for direct estimation of incidence
Exposures	Multiple exposures may be collected and examined
Outcomes	Multiple outcomes may be studied
Disadvantages	
Follow-up	Cohort studies can suffer from loss to follow-up, which results in loss of information and may introduce bias
Rare events	To study rare events in a cohort study a large number of individuals need to be followed such that it may be impractical (both in expense and logistics) to implement
Time	Cohort studies usually require a significant amount of time to acquire data to answer study questions and therefore necessitate a large commitment from the (1) investigator, (2) funding agency, and (3) participants

before the onset of clinical disease. Therefore differential recall should not occur between those who do and do not develop disease. The ability to directly estimate incidence rates for assessing the disease burden within a population is another advantage. Both designs have the advantage of being able to assess multiple types of exposures and outcomes.

The interval and clinic cohort designs also share attributes that are limitations of following individuals over time. Cohort studies always have to be concerned with loss to follow-up among participants.¹⁹ This can be due to emigration or simple unwillingness of the participant to continue with the study, although the latter issue is primarily more of a concern for interval cohort studies. This is of great concern when the rates of loss to follow-up are differential between those with and without exposure, between those who are and are not likely to develop disease, or both. Such losses would introduce bias into analyses and potentially lack validity.^{20–22} Cohort studies are challenged in studying rare events since rare diseases require a larger sample size, which may be prohibitively expensive and unfeasible. Multicenter cohort studies and cohort collaborations may be particularly useful for targeting a larger source population from which larger studies may be conducted. Finally, a cohort study often requires a considerable amount of commitment not only from the investigator, but also on from the funding agency, which may or may not be willing to commit to potentially long-term studies and in the case of interval cohort studies, from the participants who may not be willing to contribute so much of their time.

DISSIMILARITIES BETWEEN DESIGNS

In addition to the similarities between interval and clinical cohort studies, there are several differences (Fig. 1, Table 2), which may be either advantageous or disadvantageous. The advantages of one cohort design may be related to a disadvantage of the other design. Being aware of potential differences may be useful in designing a cohort study.

Potential selection biases

Recruitment issues. A distinct advantage of the interval cohort design is the ability to reach individuals that do not have access to or choose not to obtain health care. The clinical cohort is unlikely to recruit these individuals as they are not a part of the source population. This may affect the ability to generalize the study inferences to some populations.

Conversely, recruiting participants outside of health care may have disadvantages. The interval cohort design is limited to the

volunteer population that is recruited through the enrollment mechanisms utilized. Individuals willing to participate in studies may be systematically different from the general population. This could introduce volunteer and nonresponse biases.²³ For example, several studies have shown that individuals who refuse to participate have lifestyle habits that are often linked to increased disease or mortality.^{21,22,24–26} One study described nonresponders of a population-based cardiovascular study as having a higher proportion of smoking and cardiovascular disease. Furthermore, the authors stated that the participants could be classified as the “worried well” due to an overall better health status with a higher proportion of family history of cardiovascular disease than the nonparticipants.²⁴ In a different study, individuals who responded to a mailed questionnaire but refused a physical examination were more likely to smoke and have a lower education than those who participated in both a physical examination and a mailed questionnaire. Furthermore, the nonexamined men had higher total mortality rates, higher cancer mortality rates, and coronary heart disease incidence. However, these rates converged by the end of 10 years, suggesting an attenuation in the healthy participant advantage over time.²⁵

The clinical cohort also has a potential for volunteer and nonresponse bias, which may be further exacerbated when only a portion of the clinic is targeted for recruitment. However, unless individuals forgo health care, the unwillingness to participate may be attenuated, resulting in potentially lower volunteer and nonresponse bias. This would translate into an advantage for the clinical cohort design in being able to capture sicker participants within the study population. This advantage needs to be balanced by susceptibility to bias introduced by underlying referral patterns and practices of the clinic. This is analogous to a membership bias²³ whereby members of one clinic site may be different from others. Consider a hypothetical clinical cohort study in which the study population is drawn from a clinic that tends to attract the more advanced stages of disease with a higher probability of mortality. In studying this population, it is possible that there is an overrepresentation of mortality in this population, which would result in an estimate of risk that may not be transportable to the general population under care and perhaps a biased relative risk.²⁰ This referral bias can be minimized by carefully selecting a clinic that serves the majority of patients within a given geographical area such that the clinic population is then representative of all individuals undergoing care.

In evaluating a cohort study for potential selection biases, an underrepresentation (or overrepresentation) of the exposures within the study population as compared to the target population may not be of particular concern for drawing etiological associations.²⁰ Rather, the concern is whether participating in-

TABLE 2. POTENTIAL SELECTION BIASES AND WHETHER THE INTERVAL OR CLINICAL COHORT STUDY IS AT A GREATER ADVANTAGE FOR DEALING WITH THE ISSUE

<i>Interval cohort</i>	<i>Clinical cohort</i>	<i>Issue</i>
Advantage	Disadvantage	Reaching populations without health care
Disadvantage	Advantage	Subject to volunteer and nonresponse bias
Advantage	Disadvantage	Subject to membership bias
Disadvantage	Advantage	Subject to loss to follow-up among the sickest individuals
Advantage	Disadvantage	Subject to loss to follow-up among healthier individuals

dividuals may be less (or more) likely to develop disease as a result of selection biases. This would cause an underrepresentation (or overrepresentation) of the disease within the study population as compared to the target population.^{20,27} In such a situation, estimates such as the rate of disease and relative risk would be biased compared to the target population.

Retention issues. Cohort retention is an important issue in all cohort studies; whereby individuals lost to follow-up may be different from those who remain engaged in the study. For example, in one study, individuals who could not be located for follow-up interviews were more likely to have baseline characteristics that had been previously found to be predictive of higher mortality.²¹

The cohort design may influence which individuals are retained. A potential concern for interval cohort studies is that the sickest individuals may be more likely to drop out. For these individuals the burden may be too great to continue participating in the study. However, as noted earlier, it might be expected that the sickest individuals are more likely to seek clinical care and thus contribute to clinical cohorts. Individuals who remain healthy for long periods of time may potentially regard their regular clinic visits as unnecessary. These losses to follow-up would result in biased incidence rates that, if not similar in magnitude for the exposed and unexposed groups, would also bias relative measures of association. For example, in a Kaplan–Meier analysis, an interval cohort study that has losses to follow-up among individuals that were more likely to have the disease outcome would result in underrepresentation of the disease within the study population. Over time, the study population would shift toward those who are less likely to develop the disease, and thus the survival estimate would be inflated, indicating better prognosis than what would be seen in the general population. In contrast, the clinical cohort with an increased likelihood of retaining individuals who feel they need health services may result in overrepresentation of the disease and a poorer survival would be estimated than truly exists in the source population.

Potential information biases

Interval and clinical cohorts rely on different methods to collect information. The differences in visit structure will contribute to differences in the amount of information collected (Fig. 1, Table 3).

Medical records. Clinical cohorts rely heavily on patient medical records. This may be both an advantage and disadvantage.

Medical records are meant for patient care and not research. Therefore, the quality of information obtained from medical records relies on the providers' thoroughness at documenting care. Providers often overestimate the amount of information that they recorded, which may be partly due to the belief that no documentation implied that there were no problems.²⁸ Additionally, busier providers may do more than they write down, and providers who are good recorders may not be thorough in taking medical histories or giving physical examinations.^{29,30} Furthermore, the quality of information may suffer due to illegibility, missing reports, and the abstractors' interpretation and skill at finding the required information.³⁰ However, potential concerns related to the medical records may be minimized by utilizing standardized forms at patient intake. Standardized forms help reduce differing amounts of information being collected on individuals and increase the quality of data. Furthermore, with the improvement of information technology there has been a shift toward electronic medical records. Electronic medical systems are often standardized to improve clinical documentation. Thus the shift from paper to electronic medical records has improved the ability of clinical cohort studies to capture information on participants by facilitating information retrieval that is not subject to illegibility concerns.³¹

An advantage of medical records is that these documents provide precise knowledge about prescribed medications. Interval cohorts often rely on self-report of medications, which is subject to error depending on the span of time that the question in an interview addresses.³² If the study aim is to relate medications to disease outcomes, then this error may lead to nondifferential misclassification of the exposure. Nondifferential misclassification would be expected as the data are being collected prospectively on individuals in whom the outcome of interest has not occurred. Such misclassification may bias results toward the null. However, differential misclassification would occur if individuals incorrectly report medications at differing rates between those who did and did not develop disease, which may bias results either toward or away from the null similar to recall bias in case–control studies.^{23,33} However, medications are not always the exposure of interest and may be necessary information to collect as a potential confounder. Errors in the self-report of medications would then lead to residual confounding.

Medical records do not indicate whether individuals acquired the medication. Therefore clinical cohorts utilizing data from health administrative databases have an advantage of being able to access pharmacy claims and obtain precise information regarding what medications were acquired.¹⁸

TABLE 3. POTENTIAL INFORMATION BIASES AND WHETHER THE INTERVAL OR CLINICAL COHORT STUDY IS AT A GREATER ADVANTAGE FOR DEALING WITH THE ISSUE

<i>Interval cohort</i>	<i>Clinical cohort</i>	<i>Issue</i>
Advantage	Disadvantage	Information quality may vary across medical records
Disadvantage	Advantage	Medical records may provide precise information on prescribed medications
Advantage	Disadvantage	Standardized tools for data collection
Advantage	Disadvantage	Subject to medical surveillance bias
Disadvantage	Advantage	Information collected reflects actual health care needs

The interval cohort that does not primarily utilize medical records for capturing information has the advantage of having only standardized methods (e.g., interview, laboratory tests, and physical examination). However, relying on self-report for some information such as clinical diagnoses within an interview can introduce concerns about the validity of participant responses. This can be alleviated when appropriate by confirming information via medical records. It is important to note that clinical cohorts are improved when they utilize standardized methods for medical record abstraction as well as for interviews, laboratory markers, and other tools for collecting information. This helps to prevent differential ascertainment of information between exposed and unexposed groups. Furthermore, the use of registries may also overcome deficiencies in outcome ascertainment.³⁴

Visit structure. The best method to avoid medical surveillance bias in cohort studies is to ensure that the cohort undergoes systematic, standardized, and periodic data collection procedures regardless of exposures.⁶ Medical surveillance bias can be defined as a closer surveillance of individuals with seemingly medically relevant exposure for study outcomes, which may result in a higher probability of detecting disease outcomes.⁴

In studying a clinic population in which individuals may self-select or be urged toward more frequent visits by their health care provider (high-utilization individuals), it may be possible to have a higher probability for detecting disease outcomes within these individuals. Should these individuals also have an exposure of interest, then medical surveillance bias may be introduced into the study. However, the unstructured nature of the clinical cohort gives an advantage of being more likely to capture clinically relevant time points.

An unstructured visit schedule increases the complexity of the data. By having some high-utilization individuals, there is more exposure and outcome data on these individuals compared to those who seek health care services less frequently (low-utilization individuals). This introduces issues that need to be addressed in the analysis. For example, how the frequency of exposure and outcome assessment may impact the results of time-dependent analyses must be considered. Compared to a variable with a high rate of change, a variable with a slow rate of change will have lower and more similar misclassification among low and high utilization individuals. The influence upon the study results will depend on whether the misclassification is differential or nondifferential. Furthermore, when an outcome is assessed at each health-care visit, the frequency of visits may bias study results. Recently, Griffin *et al.*³⁵ showed via simulation that when biomarkers (HIV RNA and CD4 counts) are used as surrogate markers for response to HIV therapy, the measurement frequency can result in bias unless appropriate methods are used.

An interval cohort study design may not be capturing the fullest extent of the information with a set interval between study visits. The choice of interval length is informed by disease characteristics, but ultimately reflects a balance of logistical and scientific considerations. Therefore important information at biologically relevant time points may not be captured. This is especially true when dealing with biological markers that may have transient fluctuations. Therefore, an advantage

of the clinical cohort's unstructured schedule is that it may be more likely to capture important time points, as the information collected on individuals reflects actual health care.

DISCUSSION

The future for cohort studies as a fundamental tool for observational research looks bright. This is particularly true for clinical cohort studies where a variety of factors are coming together that will likely facilitate increased utilization of this design. First, as part of the Medicare Prescription Drug, Improvement, and Modernization Act of 2003, a Commission on Systemic Interoperability was established with the charge to "develop a strategy and timeline for implementing an infrastructure that supports interoperable electronic health-care systems across the nation."³⁶ This includes developing a plan for universal electronic medical records. While access and utilization of these records need to be balanced by adequate human subjects protections,³⁷ these records have the potential to serve as the core data for clinical cohort studies. However, generic electronic medical records may still lack the accuracy, level of completeness, and specificity that are required for the research setting. More work is needed to implement standardized electronic medical forms that capture data that will meet administrative and clinical care needs while providing richness for scientific research.

Second, government funding for investigator-initiated projects is now growing at a slower pace than in recent years. Reduced funds will likely lead researchers to consider utilizing information that is already "available" in medical records and electronic formats, resulting in less primary data collection, fewer laboratory tests, and the elimination of costs associated with medical personnel for clinical examinations. This would be balanced by an increase in personnel with expertise in information technology and, given the issues raised in the previous sections, personnel with expertise in epidemiological principles.

Third, there is a continuing need for improved evidence to guide clinical and health-policy decision making.³⁸ Prospective cohort studies are recognized as the leading design among non-interventional approaches³⁹ with the appropriate caveats regarding difficulties in assessing causality in observational settings.⁴⁰ Clinical cohort studies would easily complement "practical clinical trials" that have been promulgated as a way of utilizing the existing health care system to increase evidence-based knowledge.⁴¹ Methods for causal inference in observational studies have also continued to improve⁴² and are becoming increasingly adapted into clinical research papers.⁴³

While there are well-known longitudinal interval cohorts studying a diverse range of disease, why have clinical cohorts been so successful in HIV research? It may be that two aspects of HIV infection facilitate the clinical cohort design. First, HIV is a life-long infection and requires frequent contact with the health-care system for disease and therapeutic monitoring. Second, HIV infection can be monitored using widely available, minimally invasive markers of disease progression. These two factors enable individuals to be followed closely for extended period of time, facilitating the clinical cohort study design in HIV research. The paradigm for clinical cohorts, however, should be widely applicable to other disease areas.

Lastly, an emerging trend over the past several years, particularly in the HIV/AIDS field, is the collaboration among clinical cohorts to develop larger "supercohorts." Going beyond meta-analyses of published studies, these collaborative groups bring together key variables from individual clinical cohorts for addressing research questions. The EuroSIDA study is a clinical cohort study of more than 11,230 patients followed in 80 hospitals in 29 European countries and Argentina.⁴⁴ The Data Collection on Adverse Events of Anti-HIV Drugs (DAD)⁴⁵ is a study with more than 35,000 patients from 188 clinics in 21 countries in Europe, the United States, and Australia. As of February 2004, the study has accumulated more than 75,000 person-years of follow-up with the goal to assess the incidence of myocardial infarction among HIV/AIDS patients who are receiving antiretroviral therapy.⁴⁵ Seeing the success of these initiatives, the National Institute of Allergy and Infectious Diseases has launched a recent initiative entitled "International Epidemiologic Databases to Evaluate AIDS (IEDEA)" to establish a series of worldwide, regional data centers for "the compilation of data to address research questions in HIV/AIDS that are not possible to answer with currently-existing individual cohorts, to increase the generalizability of study results through use of data from different settings and populations, and to allow regions to more accurately define and monitor the HIV epidemic within their perimeters."⁴⁶

In conclusion, interval and clinical cohorts, while being able to address similar questions relating to the disease of interest, have specific advantages and disadvantages. The characteristics discussed in this paper are general in nature, as the recruitment/selection and information issues will vary by the disease, population being investigated, and the resources that are available to the investigators. Further research and evidence must be accumulated to see if the theoretical advantages and biases described here apply to real world situations.

ACKNOWLEDGMENTS

We receive funding from the National Institutes of Health [RO1-DA11602, R21-AA015032, and K24-DA00432 for the Johns Hopkins HIV Clinical Practice Cohort, U01-AI-42590 for the Women's Interagency HIV Study, and U01-AI069918 for the North American AIDS Cohort Collaboration on Research and Design, which is a part of the International Epidemiologic Databases to Evaluate AIDS (IEDEA)], and the Agency for Health Research and Quality (AHRQ-01-0012). The funding sources have had no involvement with this manuscript. We declare that we have no conflict of interest.

REFERENCES

1. Last JM, Ed.: *A Dictionary of Epidemiology*, 3rd ed. Oxford University Press, New York, NY, 1995.
2. Liddell FD: The development of cohort studies in epidemiology: A review. *J Clin Epidemiol* 1988;41:1217-1237.
3. Samet JM and Munoz A: Evolution of the cohort study. *Epidemiol Rev* 1998;20:1-14.
4. Szklo M and Nieto FJ: *Epidemiology Beyond the Basics*. Jones and Bartlett Publishers, Sudbury, MA, 2004.
5. Tager IB: Outcomes in cohort studies. *Epidemiol Rev* 1998;20:15-28.
6. Szklo M: Population-based cohort studies. *Epidemiol Rev* 1998;20:81-90.
7. Dawber TR: *The Framingham Study: The epidemiology of Atherosclerotic Disease*. Harvard University Press, Cambridge, MA, 1980.
8. Schull WJ: *Effects of Atomic Radiation: A Half-Century of Studies from Hiroshima and Nagasaki*. Wiley-Liss, New York, 1995.
9. Doll R and Hill AB: The mortality of doctors in relation to their smoking habits: A preliminary report. *Br Med J* 1954;1:1451-1455.
10. Vlahov D, Anthony JC, Munoz A, et al.: The ALIVE study: A longitudinal study of HIV infection in intravenous drug users: Description of methods. *J Drug Issues* 1991;21:759-776.
11. de Wolf F, Lange JM, Houweling JT, et al.: Numbers of CD4+ cells and the levels of core antigens of and antibodies to the human immunodeficiency virus as predictors of AIDS among seropositive homosexual men. *J Infect Dis* 1988;158:615-622.
12. Rezza G, Lazzarin A, Angarano G, et al.: The natural history of HIV infection in intravenous drug users: Risk of disease progression in a cohort of seroconverters. *AIDS* 1989;3:87-90.
13. Kaslow RA, Ostrow DG, Detels R, et al.: The Multicenter AIDS Cohort Study: Rationale, organization, and selected characteristics of the participants. *Am J Epidemiology* 1987;126:310-318.
14. Barkan SE, Melnick SL, Preston-Martin S, et al.: The Women's Interagency HIV Study. WIHS Collaborative Study Group. *Epidemiology* 1998;9:117-125.
15. Moore RD: Understanding the clinical and economic outcomes of HIV therapy: The Johns Hopkins HIV clinical practice cohort. *J Acquir Immune Defic Syndr Hum Retrovirol* 1998;17(Suppl. 1):S38-41.
16. Sudre P, Rickenbach M, Taffé P, et al.: Clinical epidemiology and research on HIV infection in Switzerland: The Swiss HIV Cohort Study 1988-2000. *Schweiz Med Wochenschr* 2000;130:1493-1500.
17. Strom B, Ed.: *Pharmacoepidemiology*, 3rd ed. John Wiley & Sons, West Sussex, UK, 2002.
18. Klein D, Hurley LB, Quesenberry CP Jr, et al.: Do protease inhibitors increase the risk for coronary heart disease in patients with HIV-1 infection? *J Acquir Immune Defic Syndr* 2002;30:471-477.
19. Hunt JR and White E: Retaining and tracking cohort study members. *Epidemiol Rev* 1998;20:57-70.
20. Criqui MH: Response bias and risk ratios in epidemiologic studies. *Am J Epidemiology* 1979;109:394-399.
21. Decoufle P, Holmgren P, Calle EE, et al.: Nonresponse and intensity of follow-up in an epidemiologic study of Vietnam-era veterans. *Am J Epidemiol* 1991;133:83-95.
22. Psaty BM, Cheadle A, Koepsell TD, et al.: Race- and ethnicity-specific characteristics of participants lost to follow-up in a telephone cohort. *Am J Epidemiol* 1994;140:161-171.
23. Sackett DL: Bias in analytic research. *J Chronic Dis* 1979;32:51-63.
24. Criqui MH, Barrett-Connor E, Austin M: Differences between respondents and non-respondents in a population-based cardiovascular disease study. *Am J Epidemiol* 1978;108:367-372.
25. Benfante R, Reed D, MacLean C, et al.: Response bias in the Honolulu Heart Program. *Am J Epidemiol* 1989;130:1088-1100.
26. Macera CA, Jackson KL, Davis DR, et al.: Patterns of non-response to a mail survey. *J Clin Epidemiol* 1990;43:1427-1430.
27. Greenland S: Response and follow-up bias in cohort studies. *Am J Epidemiol* 1977;106:184-187.
28. Osborne CE and Thompson HC: Criteria for evaluation of ambulatory child health care by chart audit: Development and testing of a methodology. Final report of the Joint Committee on Quality Assurance of Ambulatory Health Care for Children and Youth. *Pediatrics* 1975;56(4pt. 2 Suppl.):625-692.

29. Katz JN, Chang LC, Sangha O, *et al.*: Can comorbidity be measured by questionnaire rather than medical record review? *Med Care* 1996;34:73–84.
30. Luck J, Peabody JW, Dresselhaus TR, *et al.*: How well does chart abstraction measure quality? A prospective comparison of standardized patients with the medical record. *Am J Med* 2000;108:642–649.
31. Crane HM, Kadane JB, Crane PK, and Kitahata MM: Diabetes case identification methods applied to electronic medical record systems: Their use in HIV-infected patients. *Curr HIV Res* 2006;4:97–106.
32. Bodreau DM: A validation study of patient interview data and pharmacy records for antihypertensive, statin, and antidepressant medication use among older women. *Am J Epidemiol* 2004;159:308–317.
33. Werler MM, Pober BR, Nelson K, *et al.*: Reporting accuracy among mothers of malformed and nonmalformed infants. *Am J Epidemiol* 1989;129:415–421.
34. Hessol NA, Schwarcz S, Ameli N, *et al.*: Accuracy of self-reports of acquired immunodeficiency syndrome and acquired immunodeficiency syndrome-related conditions in women. *Am J Epidemiol* 2001;153:1128–1133.
35. Griffin JT, Fraser C, Gras L, *et al.*: The effect on treatment comparisons of different measurement frequencies in human immunodeficiency virus observational databases. *Am J Epidemiol* 2006;163:676–683.
36. McGee MK: Commission Ready to Work on National Health-Records System. *Information Week*. January 2005. Available from <http://www.informationweek.com/story/showArticle.jhtml?articleID=57700565>. Accessed 7 June 2005.
37. Office for Human Research Protection: Available from <http://www.hhs.gov/ohrp>. Accessed 7 June 2005.
38. Tunis SR: A clinical research strategy to support shared decision making. *Health Affairs* 2005;24:180–184.
39. Naylor CD, Guyatt GH: Users' guides to medical literature, X: How to use an article reporting variation in the outcomes of health services. *JAMA* 1996;275:554–558.
40. Holland PW: Statistics and causal inference. *J Am Stat Assoc* 1986;81:945–960.
41. Tunis SR, Stryer DB, Clancy CM: Practical clinical trials: Increasing the value of clinical research for decision making in clinical and health policy. *JAMA* 2003;290:1624–1632.
42. Little RJ and Rubin DB: Causal effects in clinical and epidemiological studies via potential outcomes: Concepts and analytical approaches. *Annu Rev Public Health* 2000;21:121–145.
43. Barron Y, Cole SR, Greenblatt RM, *et al.*: Effect of discontinuing antiretroviral therapy on survival of women initiated on highly active antiretroviral therapy. *AIDS* 2004;18:1579–1584.
44. Mocroft A, Ledergerber B, Katlama C, *et al.*: Decline in the AIDS and death rates in the EuroSIDA study: An observational study. *Lancet* 2003;362:22–29.
45. Friis-Moller N, Sabin CA, Weber R, *et al.*: Combination antiretroviral therapy and the risk of myocardial infarction. *N Engl J Med* 2003;349:1993–2003.
46. National Institutes of Health: International Epidemiologic Databases to Evaluate AIDS (IEDEA) RFA-AI-05-014. Available from <http://grants1.nih.gov/grants/guide/rfa-files/RFA-AI-05-014.html>. Accessed 7 June 2005.

Address reprint requests to:

Bryan Lau

Department of Medicine

Johns Hopkins School of Medicine

1830 E. Monument Street, Room 8070

Baltimore, Maryland 21287

E-mail: blau1@jhmi.edu