A Methodology for Conducting Retrospective Chart Review Research in Child and Adolescent Psychiatry
Robin E. Gearing PhD, MSW1; Irfan A. Mian MD, FRCPC2,3; Jim Barber PhD4; Abel Ickowicz MD, FRCPC2,3

Abstract
Introduction: Retrospective research has become largely undervalued and underutilized in child and adolescent psychiatry with the increasing singular focus on randomized control trials, despite the wealth of clinically relevant data available in historical medical records. In this paper a systematic and scientific approach to chart review research methodology for psychiatry is described. Method: Informed by available literature, a methodological stepwise approach for retrospective chart review was developed. Results: A nine step method aimed at maximizing benefits and minimizing limitations is discussed. Conclusions: Retrospective chart review is an important methodology with distinct advantages and has the potential to provide psychiatry with valuable research opportunities. This method of study should not be lost in the field of psychiatry. Key words: methodology, retrospective chart review, health record, medical record, archival data, research

Résumé

1Columbia University School of Social Work, New York, New York
2Department of Psychiatry, The Hospital for Sick Children, Toronto, Ontario
3Department of Psychiatry, Faculty of Medicine, University of Toronto, Ontario
4Deputy Vice-Chancellor, RMIT University, Melbourne, Australia
Corresponding Email: rg2372@columbia.edu
Submitted February 16, 2006; Accepted April 22, 2006

Background
Retrospective research often requires the analysis of data that were originally collected for reasons other than research (Hess, 2004; Jansen et al., 2005). This includes physician and nursing notes, ambulatory and emergency room reports, consultations, admission and discharge documentation, laboratory and diagnostic testing reports, and other clinical or administrative data.

For over eight decades, the systematic investigation of historical records has guided various clinical research (Butler & Quinlan, 1958; Wu & Ashton, 1997). The scientific utilization of existing health records is common in epidemiological investigations (Haley et al., 1980; Jansen et al., 2005), quality assessment and improvement studies (Allison et al., 2000; Kirkorian, 1979), professional education and residency training (Holmboe, Gross, & Hawkins, 1996; Neidich, 1990; Pan, Fergusson, Schweitzer., & Hebert, 2005), examination of inpatient care (Ashton, Kuykendell, Johnson, Wray, & Wu, 1995; Wu & Ashton, 1997), and in clinical research (Hellings, 2004; Rajeev, Srinath, Girimaji, Seshadri, & Singh, 2004; Staller, Kunwar, & Simionescu, 2005). Investigations using retrospective chart reviews or health record reviews have been reported to comprise 25% of all scientific articles in emergency medical journals (Worster & Haines, 2004). In comparison, while a number of psychiatric studies have successfully extracted relevant data from systematic chart reviews (Baldassano, Ghaemi, Chang, Lyman, & Lipari, 2004; Barzman et al., 2004; Bloch et al., 2005; Dworkin, 1987; Goldstein & Schwabach, 2004; Grant, 2005; Henderson et al., 2004; Marchand, Wirth, & Simon, 2004; Staller, 2004), its application is frequently limited and research findings questioned. Consequently, retrospective research is often undervalued and, hence, underutilized in psychiatry. The reluctance to use this research methodology may result from the lack of clear benefits of this approach; minimizing its recognized methodology maintains the limitations.
The advantages of conducting chart reviews include: a relatively inexpensive ability to research the rich readily accessible existing data; easier access to conditions where there is a long latency between exposure and disease, allowing the study of rare occurrences; and most importantly, the generation of hypotheses that then would be tested prospectively (Hess, 2004; VonKoss Krowchuk, Moore, & Richardson, 1995; Wu & Ashton, 1997; Worster & Haines, 2004). However, the limitations of incomplete documentation, including missing charts, information that is unrecoverable or unrecorded, difficulty interpreting information found in the documents (e.g. jargon, acronyms, photocopies, and microfiches), problematic verification of information and difficulty establishing cause and effect, variance in the quality of information recorded by medical professionals (Dworkin, 1987; Hess, 2004; Pan et al., 2005; VonKoss Krowchuk, 1995), have discouraged researchers from adopting this methodology. More developmental work is required to enhance its applicability (Wu & Ashton, 1997). We have sought to address this gap in the literature by developing a methodology for conducting retrospective chart review research in psychiatry.

**Methodological Guidelines:**

To extract data effectively and systematically from historical records requires a number of discrete steps.

**Step One: Conception** The conception stage is comprised of two components: research formulation and a clinical scan. Research formulation involves the process of articulating the research questions followed by the generation of clear hypotheses. Outlining a research question and hypothesis enables investigators to determine feasibility of retrospective chart review, instead of considering an alternative methodology. Early linking of research methodology to the study’s proposed hypothesis facilitates an informed approach that assists decisions throughout the subsequent research stages.

The second conceptual component is conducting a clinical scan of the research question and hypothesis. Seeking out clinical expertise in this stage uncovers unanticipated benefits while identifying potential methodological barriers. The earlier an investigation seeks to incorporate wider input from others, the sooner the benefits of support, expertise, sample recruitment and promotion can be achieved. Despite its sound benefits, this step often receives less attention and is rarely recognized as a definitive step (Hess, 2004).

**Step Two: Literature Review** This stage involves a systematic review of the literature pertinent to the study’s area of focus, diagnoses, conditions, demographics, criteria, and populations. A review of the literature is a standard requirement for any research initiative, including retrospective chart reviews (Findley & Daum, 1989; Hess, 2004; Jansen et al., 2005; Worster & Haines, 2004). The importance and process of conducting literature reviews is well established and it is not necessary to dwell on this step. However, it is important to note that an effective literature review requires searching several databases such as MEDLINE, PsycINFO, CINAHL, and EMBASE; it is also recommended that Boolean searches be conducted in each database. Librarians and information technologists in hospital and university libraries are valuable resources and should be consulted by investigators unfamiliar with conducting literature reviews. In addition to traditional published articles, conference proceedings and dissertations are worthy of consideration. An exhaustive literature review will provide the required background and will illuminate how other researchers operationalized key concepts or variables.

**Step Three: Proposal Development** The development of a chart review research proposal comprises writing the research proposal and operationalization of the variables. Common to all research proposals, the construction of a research proposal must include an executive summary or abstract, introduction, literature review, research question and hypotheses, methodology (design, sample, instruments, and procedure), significance of the study, limitations, budget, references, and appropriate appendices. It is also important that retrospective chart review proposals be written with some consideration for future prospective studies (Dworkin, 1987; Findley & Daum, 1989; VonKoss Krowchuk et al., 1995; Worster & Haines, 2004). Operationalization of the study variables in a review consists of two
interconnected and iterative components. First, the study variables need to be defined; these variables are generally determined by the nature and focus of the investigation. Second, study variables are then reviewed in the literature to determine how other researchers have operationalized them in similar or related investigations. It is helpful to develop an appendix comprised of concise definitions and supported with citations of studies that have similarly used each variable.

Understanding the design of existing health records and how the data is recorded is of great importance. The following strategies will assist in this process. First, it has been recommended that researchers examine the flow of information, specifically from patient to health record (Jansen et al., 2005) in order to identify established charting protocols, accepted processes of documentation, and the nature of standard documentation (e.g., emergency notes, diagnostic information, consultations, and discharge reports). Second, carefully inspect a few charts; three to five charts are recommended (Findley & Daum, 1989; Smith, 1996). This will provide critical information on how the patient chart/health record is constructed and documented. Third, consult with site-specific clinicians to ascertain how patient information is recorded and documented in multi-site studies. A clear definition of the study variables and understanding of health records provide the essential base for researchers to develop a standardized chart review data abstraction instrument.

**Step Four: Data Abstraction Instrument**
Organization, simplicity and clarity are essential criteria for the development of a uniform data abstraction instrument. Data collection should be organized in a logical order (Smith, 1996), and when possible should parallel the flow of the information in the health record. Each variable needs a simple and unambiguous response section, where the abstractors can capture the required information. Internal validity and reproducibility of any retrospective study is significantly enhanced in the standardization of the data (Jansen et al., 2005). This data abstraction instrument can be a paper or an electronic document (Allison et al., 2000). A paper instrument has some advantages, notably cost effectiveness and easier application across multiple sites. However, an electronic version is more cost effective in large investigations, reduces input error (e.g., predetermined drop down categories), and allows for easier centralization and access to data.

Finally, researchers need to decide how the data will be managed, stored, and analyzed. While pen-and-paper systems can be developed, it is strongly recommended that a data abstraction software package be used (Wu & Ashton, 1997). Several data abstraction software packages exist, such as Microsoft Access (Microsoft Corporation, Redmond, Wash) and MedQuest (Fu and Associates, Arlington, Va) (Allison et al., 2000; Banks, 1998; Worster & Haines, 2004). These packages translate the data abstraction instrument into an electronic form that can be used for data input, quality control, and the management of the data (e.g. statistical analysis and reporting). Furthermore, these programs are widely available, inexpensive, and user friendly.

**Step Five: Develop Protocols and Guidelines for Abstraction**
For any data abstraction instrument it is essential to develop a coding manual that provides a clear set of protocols and guidelines to instruct the reviewers in the collection of data (Findley & Daum, 1989; Hess, 2004; Wu & Ashton, 1997). This serves as a reference manual as to how the data will be abstracted from the health record. The manual should list each variable and explain how the variable will be captured in the data abstraction instrument, describe where the variables are located in the health record, and provide the required protocols to extract the data. Protocols with explicit criteria are designed to increase the inter-rater reliability of data abstraction (Goldman, 1992; VonKoss Krowchuk et al., 1995). Consequently, protocols generally require revisions, specifically following the piloting of the study.

**Step Six: Data Abstraction**
To abstract data effectively it is essential to understand the specific requirements of each site to determine the procedures needed to select, train, and manage the study’s data abstractors. Every health care institution has an established set of guidelines, to which all studies must adhere. While many parallels exist across institutions for conducting a retrospective chart review,
common differences among institutions are the procedures for chart procurement, retrieval rates, and access to patient charts. A number of considerations need to be taken into account, such as who may access the charts (e.g., hospital employee), the available space provided to read the charts (few institutions allow charts to be removed), the site’s hours of operation and access, and policies regarding photocopying and use of institutional or personal computers. Any of these can potentially influence effective data abstraction and require clarity before study commencement.

Data abstractors need to be carefully selected and trained (Allison et al., 2000; Pan et al., 2005; Wu & Ashton, 1997). It is preferable to select abstractors with experience in retrospective research or the area under investigation. It is also advantageous to select abstractors from health care professions, preferably with advanced levels of education (e.g., Master’s Degree). To ensure inter-rater reliability it is imperative to have a minimum of two abstractors, but it has been recommended to have four (Allison et al., 2000). It is important to determine how many abstractors are required and whether abstractors should be project based or site-specific. When possible it is preferable to use abstractors across sites rather than as on-site data collectors (Jansen et al., 2005, Jasperse & Ahmed, 1989). An alternative practice is to have one or two key abstractors train, assist, and audit site-specific data abstractors.

A standard recommendation is that data abstractors remain blind to the study hypothesis to minimize “subjectivity in classification in relation to personal theories about the study’s aims” (Worster & Haines, 2004, p. 189). Abstractors blind to the hypothesis decrease reviewer bias, specifically the possibility of their assessment being swayed by knowledge of others (e.g., investigators), concern over adversely affecting the study’s outcome, or interpreting their abstraction as too lenient or harsh (Allison et al., 2000; Chaplan, Posner, & Cheney, 1991; Goldman, 1992; Wu & Ashton, 1997). Abstractors must become familiar with a health record, be aware where the information is located, and strive to remain objective (Haley et al., 1980; Smith, 1996). Abstractors should be carefully trained with the data abstraction instrument and the accompanying protocols and guidelines. It has also been reported that the accuracy of reviewers increases when the individuals know they are being monitored (Wu & Ashton, 1997). Finally it is important to the reliability of the investigation to determine the inter-rater reliability of both the data abstraction instrument and the individual data abstractors. This can be accomplished through a pilot investigation and/or random checks.

Finally, protocols are intended to resolve any ambiguous or conflicting data. It is important for abstractors to have established management procedures to resolve any conflict. Protocols may include scheduled meetings between abstractors to resolve data conflicts, access to research investigators for clarification, or the establishment of an independent adjudication committee for consultation (Jansen et al., 2005; VonKoss Krowchuk, 1995).

Step Seven: Sample Every retrospective chart review requires a statistical power analysis to determine the appropriate sample size. Calculating the appropriate sample size is a necessary component in all research proposals and is dependant on the statistical tests used in the study. Calculating sample size is beyond the scope of this article, but can be accessed in a literature review or through consultation with a biostatistician. A rule for quickly determining sample size is 10 cases (charts) per variable, in order to obtain results that are likely to be both true and clinically useful (Sackett, Haynes, Guyatt, & Tugwell, 1991). While the literature generally holds ten events per predictor as an accepted norm (Findlay & Daum, 1989; Harrell, Lee, Muchar, & Reichert, 1985; Sackett et al., 1991), others have suggested that it is acceptable to have a minimum of seven or five events per predictor (Raykov & Wideman, 1995).

In conducting any retrospective chart review study, sampling refers to the method by which study cases or records are selected from the target population or database (Worster & Haines, 2004). Three commonly used sampling methods in retrospective chart review are convenience, quota, and systematic sampling. In convenience sampling, the most common method, suitable cases are selected over a
specific time frame; in quota sampling, a predetermined number of cases are sought from each site or diagnostic determinant; in systematic sampling, every ‘nth’ case is selected from the target population. Ascertaining the most appropriate sampling method depends on a number of factors including the importance of probability sampling, the epidemiological nature and prevalence of the specific condition, population availability, research budget, and time constraints.

The management of missing data poses methodological limitations in conducting chart review research (Hellings, 2004; VonKoss Krowchuk et al., 1995). Rules about how missing data will be handled should be devised before data collection begins (Wu & Ashton, 1997). In retrospective chart reviews, missing data can result in a hidden or non-response bias in the results, where cases with missing information may differ from the other cases (Worster & Haines, 2004). Generally, managing missing values is treated either by the deletion of the case or variable, or imputing the missing value through averaging or maximum likelihood strategies (Dworkin, 1987; Worster & Haines, 2004). In case or variable deletion, the entire case or variable is deleted from the analysis; however, this can reduce the sample size or may introduce a hidden bias (Dworkin, 1987). Imputing missing responses through statistical analysis is more common in very large computerized databases (Worster & Haines, 2004), and assumes that missing data are randomly absent. The most common maximum likelihood strategy is assigning the missing value as one response, such as with a “yes” or “no” question where the absence of a “yes” results in an immediate “no”.

There is no universal method for managing missing data, but it is imperative that researchers designing, implementing, and conducting retrospective chart review research develop protocols to address this issue. While the difficulties of missing data can often be specific to each investigation, strategies to manage this phenomenon can be garnered from anticipating common concerns associated with missing data. The most effective method to determine the development of any problems from missing data is to conduct a pilot study.

Step Eight: Ethics - It is not permissible to conduct a research study without ethics approval from an institutional review board. Therefore, it is an important step to obtain approval from the institutional review board(s) where the retrospective research will be conducted (Hess, 2004). Further, it is increasingly becoming standard that researchers conducting retrospective studies publish their ethics board approval in the methods section (Baldassano et al., 2004; Clayton & Thorne, 2000; Grant, 2005; Henderson et al., 2004; Preen, Holman, Lawrence, Baynham, & Semmens, 2004; Woogh, 2001). While the requirements for institutional review boards are standard, each board has their own protocols and policies for applicants. It is recommended that researchers contact the institution’s research ethics board coordinator, as they can provide valuable and time-saving site specific information and assistance. Any changes to the research protocols generally need to be submitted to the review board for an amended approval.

Step Nine: Pilot Study - A pilot study is essentially a small version of the proposed research. Pilot studies are critical in any study design (Perry, 2001; Van Teijlingen & Hundley, 2002). These preliminary investigations typically lack the sample size that is needed to determine statistical significance to validate a hypothesis or evaluate an instrument (Lydiard, 1991; Thompson & Spier, 1989), yet offer researchers valuable information. Specifically, pilot studies allow researchers to assess the feasibility of the planned investigation, determine the adequacy of the instrumentation, and evaluate any potential methodological pitfalls, such as data collection strategies (Prescott & Soeken, 1989). Moreover, pilot studies provide investigators with an opportunity to evaluate the reliability of their data abstraction sheet (Smith, 1996), clarify the data abstraction protocols, determine the frequency with which items are missing from the chart (Wu & Ashton, 1997), provide information on the institution’s chart retrieval rates and the process of pulling charts, and evaluate any potential sampling concerns or impact resulting from inclusion and exclusion criteria (Jansen et al., 2005).

In retrospective chart review investigations, a general recognized guideline for reliability is
that pilot studies should target ten percent of the overall sample (Gabel & Shindledeker, 1990; Wu & Ashton, 1997). In determining the actual pilot sample, the ten percent guideline needs to be adjusted to account for any concerns relating to the institution’s retrieval or availability rates and/or participant response rates if consent for their participation is required. When possible, it is preferable that pilot charts be randomly procured.

As described above, reliability is an important rationale for conducting a pilot study in this methodology. Inter-rater reliability is measured as a percentage of agreement when two or more abstractors collect data from the same chart (Allison et al., 2000). A less frequently used form of reliability is intra-rater reliability, which involves that same abstractor collecting data from the same chart on two separate occasions. Depending on the variables’ level of measurement, a Cohen’s kappa (Kappa) rating or an intra-class correlation coefficient (ICC) can statistically measure the reliability. It is recommended that a minimum of 80% and preferably 95% reliability be achieved for important variables (Allison et al., 2000; Rosen, 1995), and that variables with reliability below 70% be reassessed or re-operationalized.

Results

A comprehensive and functional nine-step method for conducting retrospective chart review research in psychiatry has been developed. See, Table 1: Methodological Steps: Conducting Retrospective Chart Review Research in Psychiatry.

The initial methodological steps begin with the conception of the retrospective research that facilitates the required literature review and subsequent proposal development. The middle steps centre on the development of a data abstraction instrument and necessary protocols and guidelines, which effectively guide the retrieval of data. The final methodological steps focus on the fundamental components of retrospective research, including data abstraction, sampling issues, ethics approval and conclude with the importance of conducting a pilot study. This nine-step model has sought to maximize benefits and minimize limitations of this methodology. Each step has highlighted components to improve the reliability and validity of this methodology, such as finding and accessing appropriate and accurate data, obtaining consistent information, improving inter-rater reliability, uniformly training data abstractors, and reducing bias. Further, this model has sought to address many of the limitations within this approach, including limiting potential abstraction and management errors, resolving ambiguous data, managing missing data, and calculating effective sample size.

Conclusion

The scientific and systematic investigation of existing health records is an important and valued methodology in health care research, specifically in epidemiology, quality assessment studies, and emergency medicine (Worster & Haines, 2004; Wu & Ashton, 1997). Notwithstanding the recognized and considerable benefits of this methodology, retrospective research has been undervalued and underutilized in psychiatry and mental health disciplines. While there remain many notable limitations to retrospective chart review research, including incomplete or missing documentation, poorly recorded, and absent information, as a methodology it continues to offer numerous advantages.

This article has offered a clear nine-step approach for conducting retrospective chart review research that may assist researchers in psychiatry and mental health to access the benefits of this methodology and minimize its limitations. Following these nine steps will increase the scientific rigour through a standard process. This process guides the clinician researcher in the process of conception and development, definition of variables, and sampling issues. The strategies outlined for procurement and abstraction of data will assist in minimizing limitations and strengthen the reliability of the data. This methodology enables clinicians to effectively conduct research that can inform and add to their practice. Retrospective chart review is an important methodology with distinct advantages and has the potential to provide us with valuable research opportunities. This method of study should not be lost in the field of psychiatry.
<table>
<thead>
<tr>
<th>Step</th>
<th>Components</th>
<th>Elements</th>
</tr>
</thead>
<tbody>
<tr>
<td>1) Conception</td>
<td>Research Formulation</td>
<td>• Design research question(s)</td>
</tr>
<tr>
<td></td>
<td>Clinical Scan</td>
<td>• Develop a hypothesis</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Use your own clinical judgment and experience</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Incorporate the clinical expertise and consultation of others</td>
</tr>
<tr>
<td>2) Literature Review</td>
<td>Literature Review</td>
<td>• Search more than one Boolean database</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Review the literature (published and unpublished studies)</td>
</tr>
<tr>
<td>3) Proposal Development</td>
<td>Write the Proposal</td>
<td>• Write the research proposal</td>
</tr>
<tr>
<td></td>
<td>Operationalize the Variables</td>
<td>• Use the chart review to plan future studies</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Define the study variables</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Examine the design of existing health records and how the data is recorded</td>
</tr>
<tr>
<td>4) Data Abstraction Instrument</td>
<td>Develop Abstraction Instrument</td>
<td>• Create a document that provides chart reviewers or data abstractors with an instrument to record the required data</td>
</tr>
<tr>
<td></td>
<td>Use Data Abstraction Software</td>
<td>• Tool can be electronic and/or paper</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Use a software package that parallels the data abstraction instrument e.g., Microsoft Access, MedQuest</td>
</tr>
<tr>
<td>5) Develop Protocols and Guidelines for Abstraction</td>
<td>Construct Coding Manual</td>
<td>• Provide a clear set of protocols and guidelines that instruct the reviewers in the collection of data (determine where and how data will be captured)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Detail rules for making decisions in ambiguous situations</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Describe how to manage missing data</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Revise as required (e.g., after pilot study)</td>
</tr>
<tr>
<td>6) Data Abstraction</td>
<td>Determine Hospital/Institutional Site Requirements</td>
<td>• Chart procurement procedures can differ across sites</td>
</tr>
<tr>
<td></td>
<td>Procedures to Select and Train Abstractors</td>
<td>• Determine site-specific retrieval rates</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Determine limits to chart access</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Selection of data abstractors (e.g., experience, profession, number, site specificity)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Training and education of data collectors</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Data abstractors remain blind to the study hypothesis</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Data abstractors must be familiar with the health records and trained in the data abstraction instrument and protocols</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Check for inter-rater reliability among abstractors</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Management of conflicting data</td>
</tr>
<tr>
<td>Step</td>
<td>Components</td>
<td>Elements</td>
</tr>
<tr>
<td>------</td>
<td>------------</td>
<td>----------</td>
</tr>
</tbody>
</table>
| 7) Sample | Sampling Issues | • Calculating sample size  
| | | • Consider sampling method  
| | | • Inclusion and exclusion criteria  
| | | • Managing missing data  
| 8) Ethics | Ethnics Review | • Obtain permission from institutional review board  
| | | • Seek review board approval for changes to the research protocols  
| 9) Pilot | Conduct Pilot Study | • Pilot studies allow researchers to assess: feasibility of the planned investigation, reliability of the data abstraction, instrumentation, effectiveness of protocols, availability of data, and sampling concerns  

**References**


Guidelines were developed for data collection from medical records for use in retrospective analyses. *Journal of Clinical Epidemiology, 58*, 268-274.


