



This document is a compilation of suggested practices for creation of a retrospective chart review form. This is not intended to be a comprehensive review of such methods, but rather a tool to assist those conducting retrospective research using chart review.

### Introduction

Retrospective analysis of medical records is a cost-effective and efficient means to collect and analyze data after an exposure has already occurred. One disadvantage of retrospective analyses is that the validity of the results depends on accurate medical records. Inconsistencies and incomplete entry in the medical record can affect the data that is abstracted during retrospective analyses. Therefore it is essential for anybody performing retrospective analyses to understand *what type of data are to be extracted and how to handle missing or incomplete records in the source documentation*.

Other than incomplete source data, the most common shortcoming of retrospective analyses is missing data resulting from lack of attention to or familiarity with data abstraction procedures. (1) The development of a well-organized retrospective review form (RRF) and RRF user manual can facilitate data abstraction, improve data quality, and increase inter-rater reliability (IRR). (2, 3)

Some best practices for creating a RRF include: a) clearly articulating the research question; b) obtaining information about the methods used to gather the data and where the data can be collected (e.g. MRI images, x-ray charts, etc.); c) identifying where the desired variable is located within the medical record; d) using a standardized RRF for the data abstraction along with a protocol that clearly outlines how the data is to be abstracted; e) performing a pilot test of the RRF and making changes as necessary; and f) continued training and mentoring for the data abstractors. (3)

### Development of RRF and Manual

For an effective RRF, the following stepwise procedures are recommended:

- In the header of each page, include the study ID and space for the data abstractor to initial and date, to indicate who collected the data and when. (4)
- List items in the order they are found in the medical record. (3)
- Group similar items together in sections, e.g. all demographics in one section, outcomes in another section, and discharge dates in a section. (4)
- Design RRF so that it is visually easy to use. Use charts and tables to help organize data; use easy to read fonts and enough spacing such that entries can be completed legibly. (4)

- Include a signature location in the footer of each page for the PI's sign-off.
- Phrase questions in a manner that will produce simple and unambiguous responses, such as yes, no, or information not available. (4)

When creating a manual for the RRF, the following should be considered:

- Outline the purpose of the study, including the rationale for data collection; information about variable names, including their definitions and synonyms; the location of the data in the medical record, if known; and general rules and guidelines for data abstraction. (3)
- Provide clear directions regarding data collection for each variable to be recorded on the RRF. (4)
- Provide an easy-to-find glossary including a list of synonyms for terms found in the medical record. (5)
- Identify which data points will require a high degree of interpretation, to be completed only by qualified (or authorized) personnel (e.g. diagnostic and surgical information) and/or reviewed and approved by such personnel. (1)
- *Provide clear instructions regarding how to treat incomplete, missing, or conflicting source data.* (2, 4)

### **Training of data abstractors and the importance of pilot studies to refined abstraction procedures:**

Training of data abstractors is essential in the successful implementation of the RRF and manual and should include the following:

- Data abstractors should be properly trained on the study protocol, the RRF, and the manual. (4, 6)
- Researchers and abstractors should be aware of the limitations and difficulties related to the creation and use of a RRF and manual. (3)
- Abstractors should be properly trained never to insert their own clinical opinions or assumptions when there is missing or incomplete data. (5)
- Regular meetings and trainings should be conducted to encourage communication, resolve conflicts, and address questions. (4)
- A pilot study should be conducted to identify weaknesses in the RRF, manual, inclusion/exclusion criteria, trainings, and survey instrument, so as to allow for modification of the RRF before the formal study begins. (2)
- During the pilot phase, the development of the RRF and manual should be considered an ongoing process, which heavily relies on the collaboration of the researchers and data abstractors. The RRF and manual should not be considered complete until reliable pilot data are collected. (3)

- Once the pilot stage ends and data collection begins, alteration of the RRF and manual should not occur. (2)
- In some cases, subsequent changes to the RRF may not be avoidable. All data abstractors must be informed of such changes to maintain uniformity in data collection. If necessary, review data that was previously collected to ensure uniformity. Changes to previously collected data should be accurately recorded with sufficient explanation, date, and initials to provide an audit trail. (3, 6)

### **Other considerations**

- Make sure all HIPAA requirements are met, including access and data security requirements.
- Retain copies of the original source data and record on the RRF where extracted data was located within the medical record. Keeping copies of source data saves time and money if you need to go back and review records.
- To the extent possible, create a database for data entry that precisely (and visually) matches the RRF, to avoid errors in data entry.
- Questionnaires sent to randomly selected patients can help to confirm the accuracy of patient records. If several inconsistencies are observed, then a questionnaire sent to all patients should be considered. If questionnaires are sent, make sure the proper protocols are completed in accordance to the IRB. (2)
- Performing random RRF and database reviews against the source data helps identify data entry errors. (2)
- Develop a data management plan to outline how extracted data will be recorded, maintained, and stored in a manner that allows for accurate reporting, interpretation, and verification. (6)
- The RRF should contain as much discrete data as possible. Variables can be consolidated or removed at a later date prior to analysis. (4)
- An advantage of a paper RRF is that it can be completed anywhere. However, use of a paper RRF requires data entry into an electronic database, which provides opportunity for error. It also requires space to store large amounts of paper.
- An electronic RRF—while requiring some initial effort up front— may decrease errors from data transposition. Electronic RRFs require abstractors to have access to an electronic device, but costs are generally not prohibitive. (4)
- Some commonly used software for abstracting clinical data are: Velos (<http://velos.com/>), RedCap (<http://www.project-redcap.org/>), or Access (<http://products.office.com/en-us/access>). Each software has pros and cons, so check with your departmental IT resource or Columbia University's Clinical Trials

Office (<http://columbiaclinicaltrials.org/>) regarding which may be of most use for you.

- An electronic RRF should include drop-down menus, error messages for incomplete fields, and space for additional notes to help increase the accuracy of data abstracted. An electronic RRF should also include a mechanism through which the Principal Investigator signs off on data collected. (5)
- Always obtain the Principal Investigator's approval of the RRF and manual and keep your PI informed throughout the data abstraction process.
- When possible, data abstractors should be blinded to the study hypothesis and/or the identity of controls, to prevent bias or interpretation of data collected from the medical record. (4)

## References

1. Pan L, Fergusson D, Schweitzer I, Hebert PC. Ensuring high accuracy of data abstracted from patient charts: the use of a standardized medical record as a training tool. *J Clin Epidemiol.* 2005;58(9):918-23.
2. Jansen ACM, van Aalst-Cohen ES, Hutten BA, Buller HR, Kastelein JJP, Prins MH. Guidelines were developed for data collection from medical records for use in retrospective analyses. *J Clin Epidemiol.* 2005;58(3):269-74.
3. Engel L, Henderson C, Fergenbaum J, Colantonio A. Medical Record Review Conduction Model for Improving Interrater Reliability of Abstracting Medical-Related Information. *Eval Health Prof.* 2009;32(3):281-98.
4. Gregory KE, Radovinsky L. Research strategies that result in optimal data collection from the patient medical record. *Appl Nurs Res.* 2012;25(2):108-16.
5. Eder C, Fullerton J, Benroth R, Lindsay SP. Pragmatic strategies that enhance the reliability of data abstracted from medical records. *Appl Nurs Res.* 2005;18(1):50-4.
6. Mihajlovic-Madzarevic V. *Clinical Trials Audit Preparation : Good Clinical Practice (GCP) Inspections.* Hoboken, NJ, USA: Wiley; 2010.