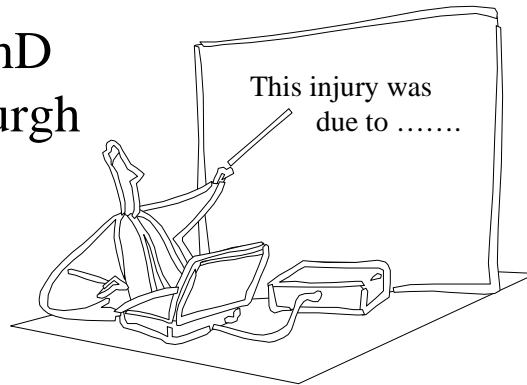


# Narrative Analysis in Injury Research

Thomas Songer, PhD  
University of Pittsburgh  
tjs@pitt.edu



Large datasets are used very frequently in injury research and injury control initiatives. This lecture introduces the concept of narrative text analysis and its use and role in injury epidemiology. Upon completing the lecture, the reader should be able to:

1. Describe the application of narrative text analysis in the injury field
2. Understand the limitations of this form of analysis

This lecture is based upon recent publications that detail the use of narrative text analysis to assess occupational injuries. For examples of narrative analysis in practice in the injury field, see the following papers:

- Jones SJ, Lyons RA. Routine narrative analysis as a screening tool to improve data quality. *Injury Prevention* 9:184-186, 2003.
- Williamson A, Feyer AM, Stout N, Driscoll T, Usher H. Use of narrative analysis for comparisons of the causes of fatal accidents in three countries; New Zealand, Australia, and the United States. *Injury Prevention* 7(Suppl 1):i15-20.
- Wellman HM, Lehto MR, Sorock GS, Smith GS. Computerized coding of injury narrative data from the National Health Interview Survey. *Accident Analysis & Prevention* 36:165-171, 2004.

# Aims

- Recognize the value of narrative data that accompanies coded data
- Provide a framework for exploring narratives
- Ultimate Goal: see to better understand circumstances of an injury

There are two primary objectives to this lecture. They include (1) having the ability to recognize the value that narrative data add to large existing datasets, and (2) understanding the foundation or framework that one can use in exploring narrative data.

Remember that the ultimate goal in this whole process is to use data to better understand the circumstances surrounding an injury.

# Qualitative Data

- Data in the form of words, language in the form of extended text
- Can be an alternative research strategy
- The underlying assumptions, methods, and research procedures of qualitative approaches are often very different from quantitative approaches

*Thomas DR*

Narrative analysis is one form of qualitative data analysis. Qualitative data are data in the form of words rather than numbers. Research designs, methods, and analytic approaches for studies that collect qualitative data are quite different from studies that rely upon quantitative data.

# Qualitative Research Approaches

- Phenomenology
- Ethnography
- Narrative analysis
- Grounded theory
- General inductive
- Discourse analysis
- Case studies

*Thomas DR*

Examples of qualitative research approaches are shown here. Narrative analysis is used frequently in the social sciences, particular in the discipline of anthropology. Sources of data in narrative analysis studies are wide ranging, but often include case studies, existing documents (such as archives and records), forms of media (newspaper, audio accounts), etc.

## Comparison of descriptive ability:

Fx ankle running on base after dark; stepped in pothole

Domain	Data Type	Nature of injury (N)	External cause (E)
<b>Medical (inpatient)</b>	ICD codes (N) STANAG (E)	824.1 Fx of ankle--Medial malleolus, open	239 Athletics/sports incl PT, other
<b>Medical (outpatient)</b>	ICD codes (N)	824.1 Fx of ankle--Medial malleolus, open	Not available
<b>Safety</b>	Hierarchical taxonomy & narrative text	Inj Type = Fracture Body Part = Foot/ankle	Activity = Running Mechanism* = Pothole Time = 2200 hrs On_Base = YES Narrative: "Jogging on running trail, detoured onto Center St due to mud . . ."

\* Derived from narrative

So what is the role of narrative analysis in injury research?

This slide illustrates why narrative data have gathered the interest of injury epidemiologists. The use of narrative data has been advocated primarily because of the existing limitations of studying injuries in medical data systems. In these systems, injuries are identified by N-codes which describe the nature of the injury (e.g. tibial fracture) and E-codes which describe the main external factor which caused the injury (e.g. motor vehicle accident). This type of classification tends to leave out a lot of the details surrounding an injury. Knowledge of the details may lead to the development of better prevention initiatives.

Consider the example illustrated here. This represents a person who got a fractured ankle from running on a military base after dark and stepping in a hole in the road. If you consider the fundamental data in the medical record to identify injuries, you will only learn that the person has a fractured ankle (N-code) due to athletics (E-code or STANAG code). This is if the person is hospitalized. If they are seen in an emergency department, it is very likely that there will not be any E-code listed. Thus, you have very limited data that detail the circumstances of the injury.

Narrative data provide more details on the injury, and in this case the narrative text identifies the activity (running), the time (2200 hours), and how the injury occurred (stepped in a hole on the road).

## Text data accompanies computerized coded data

```
029801340982110000 TWISTED ANKLE WHILE WALKING  
888015690134234520  
112952589301300000 POSSIBLE FOOD POISONING  
111436012045021439  
210001214234131980  
888121200012314521  
013298231343287821 WOUNDED BY GUNSHOT  
977237211010132123  
012128881234100010 FELL AND STRUCK BED
```

Because of the limitations of existing medical data systems, many injury epidemiologists now advocate that a short narrative text that details the circumstances of the injury should be included in medical data in addition to the information already contained within them.

Some surveillance systems are now including narrative text to supplement existing computerized data for research purposes. An example of data that appear in the National Hospital Ambulatory Medical Care Survey (NHAMCS) is shown here. Data fields exist that identify the attributes related to the medical visits. In addition, there is also a data field of narrative text included for all injuries in the NHAMCS system.

Note how this example incorporates both quantitative and qualitative data together. The primary research design, however, is quantitative in focus. It does not follow the typical methods and design issues seen in qualitative studies. This is one limitation that is inherent in narrative data added to existing data systems. This issue will be discussed shortly.

## Analytic Approaches

- Computerized keyword search using automated coding software
- manual keyword search
- manual review
  - create new variables
  - recategorize data
  - code into categories
  - identify summary themes

So, how do you analyze narrative data? No one standard exists for examining and assessing narrative data. There are, in fact, several approaches that one can take.

You can first go through the text data in varying formats. This includes reviewing each record manually (very time intensive), reviewing records with keyword searches, or buying automated coding software to conduct a keyword search (potentially expensive).

The broad goal of this review is to identify themes or common occurrences in the data. The researcher can then code or categorize these themes into quantitative data.

## Developing Categories for Analysis

- Initial categories should be determined by the objectives of the research
- Specific categories can be developed from the review of the data and the identification of frequent or significant themes, words, circumstances

*Thomas DR*

The categories that are developed for analysis are often considered from two perspectives. First, researchers should develop hypotheses based upon the research objectives of the study. These hypotheses will illustrate categories that a researcher should initially identify in reviewing the data. For example, consider the hypothesis that a specific type of equipment is involved in occupational injuries. The initial categories developed for the analysis will include this type of equipment and its variations.

The second perspective is to develop categories after reviewing the data. Categories can be developed based upon the frequency in which particular themes, words, or phrases appear in the data.



## Data Consistency

- Identify a “case definition” or “eligibility criteria”
- Go through data several times to check consistency of coding applied to the categories
- Consider the use of an independent coder

In developing categories for analysis, it is extremely important to identify a case definition or eligibility criteria to apply when reviewing the data. This will help to ensure that the data are consistent, or that the criteria that make an event a case are applied evenly throughout the study. When reviewing narrative text, researchers often find that they need to amend the eligibility criteria for a category. This is because they read of a differing circumstance that was not previously considered. Thus, it is important in this situation to review all of the records again to apply the eligibility criteria in the same fashion to each record.

## An Alternative Approach

- Analysis of narrative text structure
  - setting
  - individuals involved
  - problem
  - initiating event
  - 2nd event
  - 3rd event
  - 4th event
  - resolution

As mentioned before, there is no gold standard approach to the analysis of narrative text data. A researcher does not always have to form hypotheses first, or rely upon keyword frequencies. Another approach to starting a narrative analysis is illustrated here. This approach is taken from assessment of writing in general. In this approach, each record is reviewed to identify the listed criteria. For example, for each record, the text is reviewed to identify the setting, who was involved, the problem (e.g. injury), the event that initiated the problem, and how the problem was resolved. In this approach, the researcher is creating narrative data from narrative data.

This approach provides a framework to classifying a vast array of information. The researcher can then develop coding categories on the basis of this classification. One has to be cautious here, though, because the manner in which the original text is classified can vary from researcher to researcher. In this sense, the potential exists for the results of a narrative analysis to vary from researcher to researcher.

## Limitations

- No guidelines for what is entered as text
- Manual review is time intensive
- Purpose of data collection may not mirror underlying needs of injury research

Additional limitations exist in narrative text analysis as well, particularly with respect to its application in injury research. One of the largest concerns is in the methods applied to record narrative text. For many data systems, there are no guidelines given for what information is recorded as narrative text, and where this information is gathered from in the record. Different individuals involved in data entry may input different text. Further, the reasons why narrative text are entered into data systems may not have anything to do with injury research. In this situation, it is not possible to determine if the circumstances of the injury are not listed in the dataset because they were not in the record, or because they were in the record, but were not entered.

Another area of concern lies in the review process itself. We have previously discussed the need for consistency in the data categories developed in narrative data. However, narrative analysis is also a very time-intensive analytic procedure. Most studies are based upon manual review of narrative data. Reviewing hundreds or thousands of records and documenting data categories in this fashion requires a lot of time.

## Key Lecture Points

- Narrative analysis is another research technique in the injury field.
- It is advocated in situations where existing data contain very little information on the circumstances surrounding an injury.
- The analytic approaches in narrative analysis are often subjective, and great care must be taken to develop data consistency to support the results of the analysis.