BOOK REVIEWS

CLINICAL DATA-MINING: INTEGRATING PRACTICE AND RESEARCH. By I. Epstein, New York, NY: Oxford University Press, 2010, 228 pp. $29.95

Clinical Data-Mining (CDM), language coined by Epstein in 2001, is a research method that analyzes available agency data originally gathered for purposes other than research, to describe, evaluate, and reflect on social work practice. The purpose of the book is to facilitate a research paradigmatic change and promote Clinical Data-Mining as a practice-research method. The book is about integrating research into practice in order to encourage practitioners to be producers of knowledge, consumers of research, and implementers of research-based interventions. The book is targeted at several audiences of health and helping professions including master's degree and doctoral students.

While Epstein applied the principles of Clinical Data-Mining to the field of Social Work, the principles are equally applicable to other social fields. Clinical Data-Mining can be applied to any human service setting where there are paper, computerized, electronic, and unintended physical data that is routinely available. Data sources may include handwritten notes, word-processed or computerized records, meeting minutes, interviews, administrative spreadsheet data, standardized test scores, client questionnaires, or electronic records. Data may be stored in personal diaries, file cabinets, personal computers, and/or management information systems.

The *gold-standard* of research is randomized, controlled experiments and meta-analyses based on the experiments. While Epstein admits that Clinical Data-Mining produces less than perfect *gold-standard* research, he insists that Clinical Data-Mining generates valuable, inexpensive research knowledge for informed decision-making. Clinical Data-Mining has methodological limitation but is grounded in the logic of gold-standard research. Clinical Data-Mining offers approximations to cause-effect statements, which is the ultimate, if unreachable, objective of Clinical Data-Mining research.

The closest research methodology to Clinical Data-Mining has been *secondary analysis*, research based on quantitative databases originally amassed by researchers for research purposes. Clinical Data-Mining makes use of information that was not originally generated for research purposes. Clinical Data-Mining is an effective strategy for engaging practitioners in describing client needs, for documenting interventions, and for evaluating outcomes. Clinical Data-Mining has been successfully and productively implemented with simple to complex forms of quantitative and qualitative data analysis. Several societal changes have increased the feasibility of Clinical Data-Mining research strategies including (a) increased available information, (b) advent of agency computerization, (c) introduction of electronic records, (d) availability of personal computers, and (e) development of user-friendly, data-analytic software.

There are at least three basic methodological approaches in Clinical Data-Mining research. First, are available quantitative data that is easily converted into a quantitative database that can be statistically analyzed? Second, are data that originated in narrative form and is converted through the use of data-extraction forms from qualitative data to quantitative database and statistically analyzed? Third are available qualitative data that are converted into a database for subsequent qualitative analysis. Alternative applications of Clinical Data-Mining research strategies include combining available and original data and the application of advanced quantitative data-analysis

The term *mining* is made meaningful by a review of the steps in the Clinical Data-Mining process. The *mining* process includes:

1. Prospecting or surveying all available data sources relevant to a research question
2. Conceptualizing and organizing all independent, dependent, and intervening variables about which there are usable data
3. Deciding on the unit of analysis, whether single case, an aggregates of cases, single group, aggregates of groups, single program, aggregates of programs, single agencies, or aggregates of agencies based on the organization of data and the research question
4. Selecting a time frame in which to collect data
5. Choosing a sampling strategy and sampling size, if needed
6. Developing strategies for and forms for data-extraction
7. Securing human subjects, institutional review board, or ethics committee approval
8. Promoting validity and reliability
9. Establishing a plan of analysis, whether missing data, frequencies, percentages, central tendency measures
10. Running the univariate, bivariate, and/or multivariate analysis
11. Reviewing related literature

Qualitative Clinical Data-Mining research is rare because qualitative CDM requires unusually rich deposits of available data and data of the richness required are rarely available for retrospective analysis. A second reason for the rarity of qualitative CDM is that researchers have not thought about applying a CDM approach in research. Epstein admits that CDM qualitative research is quite similar to conventional qualitative research, except in a CDM approach available rather than original material is used.

There are pitfalls in the application of Clinical Data-Mining research strategies. CDM is often labor-intensive, tedious, and occasionally dirty work. A CDM strategy does not lend itself to an application for large research grants. Oftentimes, the variables that are of greatest interest are not available. Problems of validity and reliability confront all CDM research. CDM has the potential of surfacing political and organizational conflicts in bureaucratic settings.

There are many available databases in the fields of psychology, sociology, education, and social work that could be accessed for research purposes. For example, the OGS admissions application has information that could be appropriated for research purposes including degree program choice, entry (and later exit) date, geographic location, language, citizenship, birth date,

birth place, gender, racial category, previous colleges attended, and employment. Religious data will usually require original research that can be joined with available data for a focus on the integration of religion and society.

A database that is rich in sociological information is the online United States census. The census includes such information as (a) age, gender, racial, and household issues; (b) language and educational backgrounds; and (c) housing arrangements, business profiles, and population trends and density. The information in the census can be analyzed or used as baseline data that is compared to original data.

Some years ago a private college in western Kentucky applied a CDM approach in building a case for gaining permission to offer a degree completion program in southern Illinois. The available database was the online United States Census. The census gave the number of college graduates for each county in Illinois in real numbers and percentages. The case presented to the high education authority in Illinois was the percentage of residents with bachelor degrees in a southern Illinois community college district compared to the percentages for the state of Illinois. The comparison showed that the community college district was significantly underserved for bachelor degrees. With 37 applications presented to the Illinois higher education authority, four out-of-state agencies were given permission to offer college programs in Illinois. The application of the private college in western Kentucky was approved and then some. Instead of awarding permission to offer the degree completion program in one community college district in southern Illinois, permission was given to offer degree completion college courses in three community college districts. The case was built entirely on an available database that may not have been primarily developed for research purposes. Clinical Data-Mining has potential for meaningful research by graduate students.

L. Standridge
Oxford Graduate School