Craig Alder

Dr. Higgins and Dr. Hertzberger

NRSG 741

8 February 2017

Assignment 2, Critique 1

Samuels, et al. (2015) present a summary of the current state of healthcare data available

for nurse practitioners and researchers, including opportunities and challenges. The authors

explain that because of changes in electronic record keeping brought about by the Health

Information Technology for Economic and Clinical Health Act, there is an unprecedented amount of

electronic health record (EHR) data available for nursing research. They cite a Robert Wood

Johnson report showing 44% basic EHR use by hospitals. This is important because there is more

data available that can be utilized, challenges for data security, different standardization of EHR

data across institutions (and within), complex possibilities for technical analysis, new bureaucratic

regulations, and greater overall opportunity for efficiency. Accordingly, the authors argue that

quality assurance and full utilization of data must all be addressed.

The authors spend substantive time on the difference between structured and unstructured

documentation. Structured fields are discrete, quantitative data points. Unstructured fields are

qualitative notes, which are more cumbersome for data analysis. The authors cite research

suggesting up to 80% of hospital data is unstructured. Statistical analysis called Sentiment Analysis

allows bodies of text to be sorted categorically with 80% accuracy, but this is time consuming and

usually means less total sample data. The authors cite further research showing that at one

research project collected 4,500 points of data on pain management compared to 810,774 data

points on pain analyzed from a U.K. hospital in a separate study. I would question the validity of

comparing these two studies, especially without further context regarding the overall research

designs and goals of each study. Samuels, et al. (2015) point out that of concern to both structured and unstructured data, this approach to healthcare analytics philosophically appears positivistic, simplifying complex questions to mere math.

The authors go on to elucidate Big data -- characterized by mass, speed, and range of different data -- as allowing remarkably accurate descriptions of what is occurring in the social world, though not necessarily explanations of why. Data mining utilizes software to identify patterns and relationships in these large data sets. Statistical programs identify the most salient information in a data set, replacing the tedious, small-sample approach of hypothesis testing.

The authors argue that in order to make all this more effective, data collection and operationalization of variables needs to work towards standardization, privacy concerns must be efficiently addressed, and analytical techniques for unstructured data needs to be refined. As a review of previous research, the authors do not present any original data analysis to support their descriptions. Most of the claims cited to support their descriptions do not include more than the conclusion made by the other researchers, but the claims are logically consistent. Some claims, however, do seem to lack evidence. Specifically, the authors claim that structured fields of data allow access to 100% of "the population." They argue, without evidence, that nursing has the potential to function like large companies that custom-prepare ads based on big data. Also, they argue that current organizations and government initiatives are important for creating appropriate systems. Evidence supporting these claims would bolster what is already a clear review with specific implications reviewed above. One of the final critiques I would make is that big data may include risks institutionalizing biases and prejudices, whether from machine learning that does not no better or researchers creating seemingly innocuous procedures with ramifications as big as the data being analyzed.