Scope of Informatics Problem in Research Statistics

Ryan Gosling, Emma Watson, Ban Ki-Moon

University of Hawaii

Scope of Informatics Problem in Research Statistics

According to the Centers for Disease Control and Prevention (CDC), sexually transmitted diseases (STDs) remain a major public health challenge in the United States (CDC, 2010). CDC estimates 19 million new STD cases each year, costing the healthcare system $16.4 billion annually, and resulting in long term health consequences for infected individuals (CDC, 2010). National profile shows trends by age, gender, race/ethnicity, and location in the United States (CDC, 2010). Using a public database, our group is investigating the prevalence of STDs among gender. Lange & Jacox (1993) stated, “Large clinical data bases can be used to develop risk profiles and then to identify patients who match the profile so that preventive action can be taken” (p. 54). Healthcare professionals have a unique opportunity to provide education and counseling to the at-risk population as primary prevention is the key in reducing national STDs (CDC, 2010). This paper will introduce the scope of an informatics problem, dataset that contains data to answer a clinical question, data structure, the clinical question, the results, and the evaluation of the data set’s usefulness.

**Formulation and Scope of Informatics Problem**

Healthcare is currently facing a challenge with the complexity of a large amount of dynamic data coming from different sources that make up the patient health record (UA NURS 646, 2011). In order to provide efficient and effective patient care, it is imperative healthcare informatics aid in allowing healthcare providers information access, communication among team members, and standardization of practice (Bakken, 2004). Currently information is collected by an assortment of different government and private agencies, making information sharing and translation difficult (Zeni & Kogan, 2007).

**Dataset and Source**

To answer the clinical question, a dataset is retrieved from the National Center for Health Statistics (NCHS) website. NCHS is the nation’s principal health statistics agency and conducts a variety of population-based studies to assess and monitor the health of the nation (Zeni & Kogan, 2007). According to the NCHS website, the agency collects data from “birth and death records, medical records, interview surveys, and through direct physical exams and laboratory testing” (CDC, 2007). Due to the website reputability, the data was accessed via the Wide-ranging Online Data for Epidemiologic Research (WONDER) (CDC, 2007).

**Structure of the Data**

The dataset is retrieved as a raw data file containing columns, separated by a comma. Columns include the year of the study, ethnicity, age, gender, and reported STD of the participants. Ethnicity, age, gender, and reported STD are nominal data while the year of the study is interval ratio level of measurement. With large amounts of data available from multiple sources it is important to have a structured dataset that has standardization of nomenclature when capturing the data to provide the capability of easily storing and retrieving data for analysis (Zielstorff, 1998). Standardization allows the data to be consistently shared and reused, for the information to be comprehended by both computers and humans (Dykes et al., 2009).

**Question Posed**

Are STDs more prevalent among females compared to males? The dataset contains information that will be used to answer this question, which could have been answered with an informatics system. When describing an informatics system, Coiera (2003, p. 115) states, “The computer provides mechanisms for capturing information during the clinical encounter, stores it in some secure fashion, and permits retrieval of that information by those with a clinical need.”

**Results**

The dataset includes a sample of a population comprising 1,503 males and 1,494 females infected with STDs from 1996 through 2009. The study focused on finding if significant differences exist in STDs infection among gender. Nominal data, including gender and STDs were sorted into a pivot table and a Chi-square method was applied. The results yielded a chi-square = 0.06, and a p = 0.97. It can be concluded that there is not a statistically significant difference between males and females in having STDs.

**Evaluation of the data set’s usefulness**

The dataset proved to be useful in answering the clinical question posed, based on its ability to be generalized to larger samples within the population. The dataset was freely downloaded, as ASCII format, which is easily imported in different statistical tools. This study has demonstrated how large public databases are vital in research studies to answer nursing clinical questions. However, the study also demonstrated that navigating various federal government agencies is time consuming (Zeni & Kogan, 2007), which does not always fit the busy nursing workload. Nurses need knowledge of how to access the databases, how to do the statistical analysis and interpret them (lange & Jacox, 1993). Also, the clinical question is limited to the data included in the database (Zeni & Kogan, 2007). “Nurses need to seek knowledge to support their practice as a routine part of their clinical work, just as much as ordering a laboratory test or conducting a patient physical examination” (Coiera, 2003, p. 65). The informatics system is imperative for nursing decision-making and patient safety, as McGonigle & Mastrian (2009, p. 16) states, “there is need to remedy healthcare information technology concerns, challenges, and issues that we face today.”

References

Bakken, S., Cimino, J. J., & Hripcsak, G. (2004). Promoting Patient Safety and Enabling Evidence-Based Practice Through Informatics Medical Care, 42, Number 2 Supplement, II-49 - II-56

Centers For Disease Control And Prevention 20110607 CDC WonderCenters for Disease Control and Prevention (2011, June 7). CDC Wonder. Retrieved October 19, 2011, from <http://wonder.cdc.gov>

Centers For Disease Control And Prevention 20110607 CDC WonderCenters for Disease Control and Prevention (2010). CDC Wonder. Retrieved October 19, 2011, from http://www.cdc.gov/std/stats09/trends.htm

Coiera, E. (2003). Guide to Health Informatics (2nd ed.). London: Arnold Publishers.

Dykes, P., Hyeon-Eui, K., Goldsmith, D., Choi, J., Esumi, K, & Goldberg, H. (2009). [The adequacy of ICNP Version 1.0 as a representation model for electronic nursing assessment documentation.](https://nursing.d2l.arizona.edu/content/nursing/FA11/NURS646/PDF%20Files/Week8/The%20adequacy%20of%20ICNP%20Version%201.0%20as%20a%20representation%20model%20for%20electronic%20nursing%20assessment%20documentation.pdf?_&d2lSessionVal=5C8YxJrqNd3RsW1kpIDJZ7iYI) JAMIA 16, 238-246

Heavey, E. (2011). Statistics for nursing: A practical approach. Sudbury, MA: Jones & Bartlett.

Lange, L. L., & Jacox, A. (1993). [Using large databases in nursing and health policy](https://nursing.d2l.arizona.edu/content/nursing/FA11/NURS646/PDF%20Files/Usinglargedatasets.pdf?_&d2lSessionVal=9eKHSy3UBGcmhApt8Dn0k5hLj), *Journal of Professional Nursing, 9*(4), 204-211

McGonigle, D. & Mastrian, K. (2009). Nursing Informatics and the Foundation of Knowledge (2nd ed.). Jones & Bartlett.

Minarik, P. A., & Bozzo, J. (1999). Using hospital databases. *The American Journal of Nursing, 99*(2), 54

University of Arizona (UA) NURS 646 (2011). Mini-lecture: Healthcare Informatics: Theory & Practice.Tucson, AZ

US Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for HIV, STD and TB Prevention (NCHSTP), Division of STD/HIV Prevention, Sexually Transmitted Disease Morbidity for selected STDs by age, race/ethnicity and gender 1996-2009, CDC WONDER On-line Database, June 2011. Accessed at http://wonder.cdc.gov/std-v2009-race-age.html on Oct 16, 2011 8:26:33 PM doi:10.1016/j.outlook.2006.09.007

Zeni, M. B., & Kogan, M. D. (2007). Existing population-based health databases: Useful resources for nursing research. *Nursing Outlook, 55*(1), 20-30

Zielstorff, R., (1998). Characteristics of a good nursing nomenclature from an informatics perspective. *Online Journal of Issues in Nursing, 3, (2)*