



Summer Institute in Nursing Informatics 2019  
Poster Presentation

## **Evaluation and improvement of the Data Quality Toolkit: Supporting Collaborative Data Quality Improvement in Healthcare Organizations**

Urmita Banerjee; Yili Zhang, PhD; and Gunes Koru, PhD

**Introduction:** Data quality problems, which are prevalent in healthcare organizations, often prevent data analytics efforts from reaching their full potentials. Data Quality Toolkit (DQT) is a multi-user client-server software application which supports collaborative data quality improvement involving healthcare professionals with different roles, background, and knowledge. Following the development of DQT in our lab, it was essential to evaluate its effectiveness in a real setting and further improve it before its pilot implementation in a state health agency. Accordingly, this research had three objectives: 1) Evaluating the perceived usefulness of DQT and identifying directions for improvement, 2) Evaluating the perceived ease of use of DQT and identifying directions for improvement, and 3) Improving DQT according to the identified directions. **Methods:** This study adopted a qualitative approach to obtain rich contextual data for evaluation and improvement of DQT. The opinions, suggestions, and experiences related to working with DQT in terms of its usefulness and usability were collected through focus groups conducted with graduate student participants. Focus groups allowed eliciting both individual expressions and group perspectives. Prior to focus groups, the participants used DQT and conducted month-long data quality improvement projects in the tool within their assigned groups by detecting defects, tracking them and resolving them. Each group worked on a different simulated home care readmission dataset injected with data defects. The focus groups were audio recorded and notes were taken. The transcripts of the recordings were produced and analyzed by using the Framework method along with the notes. To ensure the validity of evidence, triangulation was performed by collecting and analyzing information provided through direct opinions submitted by groups in a report. According to analysis results, DQT was improved by introducing new features and fixing the programming errors. **Results:** Capabilities of DQT such as establishing and managing metadata about business rules that apply to healthcare data as constraints, efficient and fast defect detection and data and business rule versioning were perceived to be useful. Suggested improvements centered around strengthening communication and collaborative work. The participants mentioned a need to enter metadata describing the original data and be able to track user, data and constraints inter dependency in finding defects. Regarding usability, accurate defect detection reports and the ease of operation in identifying different defect kinds in the dataset was perceived to improve ease of use. For usability improvement, simplifying task navigation and layout and increasing the overall intuitiveness and appeal of the interface were suggested. Accordingly, the interfaces were improved to address the usability issues, and new features were introduced to DQT that allow viewing dependencies on other team members over the data elements and constraints and enable maintaining a data dictionary. **Conclusion:** This research demonstrated that the participants were indeed able to use DQT for collaborative data quality improvement, which is highly promising. Their feedback helped improve the usefulness and usability of DQT. As the next step, we plan to implement DQT as a pilot in the state health agency for further improvements and refinements.