INTERNATIONAL JOURNAL OF MEDICAL INFORMATICS 84 (2015) 248–262

journal homepage: www.ijmijournal.com

A multiple-scenario assessment of the effect of a continuous-care, guideline-based decision support system on clinicians' compliance to clinical guidelines

Erez Shalom[a,*], *Yuval Shahar*[a], *Yisrael Parmet*[b], *Eitan Lunenfeld*[c]

[a] *Medical Informatics Research Center, Department of Information Systems Engineering, Ben-Gurion University of the Negev, Beer-Sheva, Israel*
[b] *Department of Industrial Engineering, Ben-Gurion University of the Negev, Beer-Sheva, Israel*
[c] *Department of Obstetrics and Gynecology, Soroka Medical Center, Faculty of Health Sciences, Ben-Gurion University of the Negev, Beer-Sheva, Israel*

ARTICLE INFO

Article history:
Received 3 September 2014
Received in revised form
8 January 2015
Accepted 9 January 2015

Keywords:
Medical informatics
Clinical decision support systems
Therapy
Computer-assisted
protocol-directed
Clinical guidelines
Quantitative evaluation
Preeclampsia/eclampsia

ABSTRACT

Objectives: To quantify the effect of a new continuous-care guideline (GL)-application engine, the Picard decision support system (DSS) engine, on the correctness and completeness of clinicians' decisions relative to an established clinical GL, and to assess the clinicians' attitudes towards a specific DSS.

Methods: Thirty-six clinicians, including residents at different training levels and board-certified specialists at an academic OB/GYN department that handles around 15,000 deliveries annually, agreed to evaluate our continuous-care guideline-based DSS and to perform a cross-over assessment of the effects of using our guideline-based DSS. We generated electronic patient records that realistically simulated the longitudinal course of six different clinical scenarios of the preeclampsia/eclampsia/toxemia (PET) GL, encompassing 60 different decision points in total. Each clinician managed three scenarios manually without the Picard DSS engine (Non-DSS mode) and three scenarios when assisted by the Picard DSS engine (DSS mode). The main measures in both modes were correctness and completeness of actions relative to the PET GL. Correctness was further decomposed into necessary and redundant actions, relative to the guideline and the actual patient data. At the end of the assessment, a questionnaire was administered to the clinicians to assess their perceptions regarding use of the DSS.

Results: With respect to completeness, the clinicians applied approximately 41% of the GL's recommended actions in the non-DSS mode. Completeness increased to the performance of approximately 93% of the guideline's recommended actions, when using the DSS mode. With respect to correctness, approximately 94.5% of the clinicians' decisions in the non-DSS mode were correct. However, these included 68% of the actions that were correct but redundant, given the patient's data (e.g., repeating tests that had been performed), and 27% of the actions, which were necessary in the context of the GL and of the given scenario. Only

* *Corresponding author*. Tel.: +97286477160.
E-mail address: erezsh@bgu.ac.il (E. Shalom).

http://dx.doi.org/10.1016/j.ijmedinf.2015.01.004

5.5% of the decisions were definite errors. In the DSS mode, 94% of the clinicians' decisions were correct, which included 3% that were correct but redundant, and 91% of the actions that were correct and necessary in the context of the GL and of the given scenario. Only 6% of the DSS-mode decisions were erroneous. The DSS was assessed by the clinicians as potentially useful.

Discussion: Support from the GL-based DSS led to uniformity in the quality of the decisions, regardless of the particular clinician, any particular clinical scenario, any particular decision point, or any decision type within the scenarios. Using the DSS dramatically enhances completeness (i.e., performance of guideline-based recommendations) and seems to prevent the performance of most of the redundant actions, but does not seem to affect the rate of performance of incorrect actions. The redundancy rate is enhanced by similar recent findings in recent studies. Clinicians mostly find this support to be potentially useful for their daily practice.

Conclusion: A continuous-care GL-based DSS, such as the Picard DSS engine, has the potential to prevent most errors of omission by ensuring uniformly high quality of clinical decision making (relative to a GL-based norm), due to the increased adherence (i.e., completeness) to the GL, and most of the errors of commission that increase therapy costs, by reducing the rate of redundant actions. However, to prevent clinical errors of commission, the DSS needs to be accompanied by additional modules, such as automated control of the quality of the physician's actual actions.

1. Introduction

1.1. *The need for an evaluation of the value of automated support to guideline-based care*

Clinical Guidelines (GLs) are one of the manifestations of the recent emphasis on *evidence-based medicine*, which tries to apply the best available evidence gained from scientific methods such as research studies, meta-analyses, and reviews, to support better clinical decision making [1]. Extensive evidence confirms that state-of-the-art GLs are a powerful method for standardization and uniform improvement of the quality of medical care and patient outcomes, often increasing patient survival rates while reducing the escalating costs of medical care [2–6].

Despite these useful findings, the level of adherence to GLs in daily practice is relatively low [2,3,7]. Therefore, care providers, health-care managers, and patients would benefit from automated support of GL-based care through the use of computerized GL-based decision support systems (DSSs). These systems include an electronic representation of GLs and support their automated dissemination and application at the point of care [8–11]. Over the past two decades, there have been a number of efforts to support the application of complex GLs in an automated fashion, typically providing static, one-time recommendations at several distinct points along the process of care [12–15]. However, none of these frameworks fully supports a continuous application of the GLs over significant stretches of time, providing recommendations when necessary, handling issues such as missing data in the electronic medical record (EMR), and also supporting a data-driven, asynchronous application (i.e., not just during a session with the care provider). Furthermore, there are very few large-scale assessments of the potential effect of using a GL-based DSS on the continuous application of a complex GL over time, especially assessments that use a meaningful number of clinicians. According to Isern [11] and others [15–17], there is a relative lack of research on the effects of GL application on the quality of clinical decisions by clinicians (i.e., their level of adherence to the GL's recommendations, and the percentage of their decisions that is correct according to the GL) and of "*in vivo*" evaluations in the area of GL application engines.

In a recent comprehensive methodological review summarizing the past decade's research regarding the life cycle of computerized GLs [18], Peleg noted that in general, only very few evaluations of GL-based DSSs have been made, since a full evaluation is often complicated by the fact that the DSS might allow clinicians to deviate from the GL's recommendations. Like others [11,15–17], Peleg concluded that additional research should be performed on the effect of GL-based DSSs on clinicians' behavior, in particular on improving their compliance to GLs.

1.2. *The objectives of this study*

The main objective of this study was to quantitatively evaluate the effects of a longitudinal GL-based DSS framework, designed for realistic, continuous use over multiple sessions, on the quality of medical decisions made by a group of physicians. We previously designed and implemented such a continuous-care framework (see Section 2.1). In the current study, we used a set of realistically simulated longitudinal medical records of patients, each presenting one of multiple clinical, GL-based scenarios that need management according to a known, well-established obstetrics GL for management of preeclampsia/eclampsia.

As a secondary objective, we wanted to assess the subjective perception of the clinicians regarding the GL-based decision-support framework.