Journal of Biomedical Informatics 57 (2015) 288–307

Contents lists available at ScienceDirect

Journal of Biomedical Informatics

journal homepage: www.elsevier.com/locate/yjbin

Hypothesis generation using network structures on community health center cancer-screening performance

Timothy Jay Carney [a,g,h,*], Geoffrey P. Morgan [b], Josette Jones [a], Anna M. McDaniel [a,c], Michael T. Weaver [c], Bryan Weiner [g], David A. Haggstrom [d,e,f]

[a] Indiana University School of Informatics (IUPUI), United States
[b] Carnegie Mellon University, Computational Analysis of Social and Organizational Systems (CASOS), United States
[c] Indiana University School of Nursing, United States
[d] VA HSR&D Center of Excellence on Implementing Evidence-based Practice, Roudebush VA Medical Center, United States
[e] Division of General Internal Medicine, Indiana University School of Medicine, United States
[f] IU Center for Health Services and Outcomes Research, Regenstrief Institute, Inc., United States
[g] University of North Carolina, Gillings School of Global Public Health, United States
[h] Global Health Equity Intelligence Collaborative (GHEIC), LLC, United States

ARTICLE INFO

Article history:
Received 30 September 2014
Revised 31 July 2015
Accepted 4 August 2015
Available online 11 August 2015

Keywords:
Computational modeling
Simulation
Community health centers
Systems-thinking
Cancer screening
Network theory
Learning health system
Health Disparities

ABSTRACT

Research objectives: Nationally sponsored cancer-care quality-improvement efforts have been deployed in community health centers to increase breast, cervical, and colorectal cancer-screening rates among vulnerable populations. Despite several immediate and short-term gains, screening rates remain below national benchmark objectives. Overall improvement has been both difficult to sustain over time in some organizational settings and/or challenging to diffuse to other settings as repeatable best practices. Reasons for this include facility-level changes, which typically occur in dynamic organizational environments that are complex, adaptive, and unpredictable. This study seeks to understand the factors that shape community health center facility-level cancer-screening performance over time. This study applies a computational-modeling approach, combining principles of health-services research, health informatics, network theory, and systems science. *Methods:* To investigate the roles of knowledge acquisition, retention, and sharing within the setting of the community health center and to examine their effects on the relationship between clinical decision support capabilities and improvement in cancer-screening rate improvement, we employed Construct-TM to create simulated community health centers using previously collected point-in-time survey data. Construct-TM is a multi-agent model of network evolution. Because social, knowledge, and belief networks co-evolve, groups and organizations are treated as complex systems to capture the variability of human and organizational factors. In Construct-TM, individuals and groups interact by communicating, learning, and making decisions in a continuous cycle. Data from the survey was used to differentiate high-performing simulated community health centers from low-performing ones based on computer-based decision support usage and self-reported cancer-screening improvement. *Results:* This virtual experiment revealed that patterns of overall network symmetry, agent cohesion, and connectedness varied by community health center performance level. Visual assessment of both the agent-to-agent knowledge sharing network and agent-to-resource knowledge use network diagrams demonstrated that community health centers labeled as high performers typically showed higher levels of collaboration and cohesiveness among agent classes, faster knowledge-absorption rates, and fewer agents that were unconnected to key knowledge resources. *Conclusions and research implications:* Using the point-in-time survey data outlining community health center cancer-screening practices, our computational model successfully distinguished between high and low performers. Results indicated that high-performance environments displayed distinctive network characteristics in patterns of interaction among agents, as well as in the access and utilization of key knowledge resources. Our study demonstrated how non-network-specific data obtained from a point-in-time survey can be employed to forecast community health center performance over time, thereby enhancing the sustainability of long-term strategic-improvement efforts. Our results revealed a strategic profile for community health center cancer-screening improvement via simulation over a projected 10-year period.

* Corresponding author at: University of North Carolina, Gillings School of Global Public Health, Chapel Hill, NC, United States.
E-mail addresses: tcarney@unc.edu (T.J. Carney), gmorgan@cs.cmu.edu (G.P. Morgan), jofjones@iupui.edu (J. Jones), amcdanie@iupui.edu (A.M. McDaniel), mtweaver@iupui.edu (M.T. Weaver), weiner@email.unc.edu (B. Weiner), dahaggst@iupui.edu (D.A. Haggstrom).

http://dx.doi.org/10.1016/j.jbi.2015.08.005
1532-0464/

The use of computational modeling allows additional inferential knowledge to be drawn from existing data when examining organizational performance in increasingly complex environments.

1. Introduction

Improving cancer-screening performance for breast, cervical, and colorectal cancer in community health centers (CHCs) is a priority [1]. Cancer screening rates among vulnerable populations typically served by CHCs remain below the nationally targeted benchmarks [2,3]. Low cancer-screening rates are primary contributors to cancer health disparities among this population, resulting in an increase in the number of new cancer cases, increased mortality and lower five-year survival rates [1,3,4].

The Health Disparities Cancer Collaborative (HDCC) established in 2003–2005 represents a structured approach toward building capacity, encouraging best practices, and evaluating the areas of deficiency in cancer-care delivery as it contributes to present and future cancer-screening performance levels [1,3]. The HDCC, co-sponsored by the Health Resources Services Administration (HRSA) and the National Cancer Institute (NCI), includes CHCs from around the country [5].

Despite advancements in facility-level cancer-screening rates among HDCC participants, two major performance issues regarding the sustainability of effort over time and diffusion of best practices have emerged. Previous studies have revealed that HDCC participation was positively correlated with improvements in screening for breast, cervical, and colorectal cancer with improvements derived through providers' self-reported measures over the previous year [1,3]. Other studies also revealed that maintaining improvements in process outcomes well after their HDCC participation remains a major challenge for CHCs [6–9]. Additionally, the best practices that discriminate high-performing CHCs from low-performing CHCs are not easily duplicated in low-performing CHCs.

Issues related to the sustainability and diffusion of innovation reveal the possibility of additional organizational and/or practice-setting factors that could affect health outcomes. Such organizational factors among HDCC participants may not be easily decipherable or explained through the use of traditional statistical modeling. In an earlier study, we used traditional (statistical) modeling to examine the correlation between antecedents and outcome variables, as collected from a single point-in-time snapshot of CHCs cancer-screening practices. Such correlations, while at times positive, may not be reliable necessarily for the accurate prediction of future organizational practices and/or outcomes.

We address this issue by a hypothesis-generating experiment of CHCs cancer-screening practices, using dynamic network-simulation analysis to convert single point-in-time survey data into a dynamic network-analysis data source and to generate a series of network diagrams (configurations) to be used to compare high-performing CHCs to low-performing CHCs.

2. Using computational modeling to evaluate community health centers' practices

Computational organizational models (COMs) are useful in situations where actual experimentation on the population of interest is not feasible or is deemed unethical. Such scenarios may be actual (having already occurred in past) or hypothetical (providing an interesting future possibility). In this project, we used an existing COM called Construct-TM, which has been validated previously to reflect the dynamics of group diffusion of information accurately [10].

Recent advances in social networks, cognitive sciences, computer science, and organizational theory have led to a new perspectives on organizations, accounting for both their computational nature and their underlying social and knowledge networks [11]. Organizations are complex, computational, and adaptive agents in their own right [12], as they are composed of other elements which are constrained and enabled by their positions in social settings and knowledge webs of affiliations, linked agents, and tasks. Computational modeling allows for in-depth investigations between an organization's complex and adaptive nature that may include, (1) the interaction of the specific agents/actors, (2) the resources present in the organization (e.g. the use of information-technology, clinical reminders, prompts at point-of-care, etc.), and (3) the core activities that directly or indirectly impact the desired health objectives and outcomes [13,14].

Computational modeling enables the construction of a virtual model for a system, such as a hospital or patient-care unit, which can be used to study its behavior under various conditions [15,16] as well as to generate hypotheses regarding organizational dynamics [17,18]. Traditional statistics cannot explore adequately the *what-if* scenarios and are unable to investigate hidden relationships between people and resource configurations, which are necessary to explain the organizational behaviors and/or outcomes of which computational modeling is capable [19,15]. Computational models can provide meaningful insights into organizational behavior (e.g. linking a set of organizational predictors of outcomes to observable patterns of key resource-utilization). The use of computational models, specifically simulation, allows for the generation and testing of hypotheses [17]. Computational models used in conjunction with hypothesis testing could be an effective approach to resolving complex organizational challenges.

The goal of this research is to explore the possible existence of simple, nonlinear processes underlying team or group behavior [17] through computational models. Our core objectives are to determine if we can (1) duplicate CHC performance over an extended period of time into the future using simulations, which could be used for the formulation of hypotheses on sustainability; (2) identify structural differences between observed high-performing CHCs and low-performing CHCs to generate hypotheses on issues related to the diffusion of best practices; and (3) examine and analyze point-in-time survey data, such as survey data collected in 2006 on HDCC breast, cervical, and colorectal screening practices, and determine if exploratory computational analysis can add value to existing information.

In this study, we focus on the use of existing simulation tools to examine the dynamics of CHCs. We discuss Construct-TM in the next section, followed by a discussion of the use of NCI/HRSA HDCC survey items to define the Construct-TM model.

2.1. Construct-TM overview

Construct-TM is "a social network simulator" [20] based on the concept of transactive memory (the "TM" in Construct-TM), which is the process by which a group of people (e.g. an organization such as a community health center) collectively store, retrieve, learn, communicate (both inside and outside the group), and use knowledge [20]. Construct-TM employs dynamic-network theory to