HHS Public Access
Author manuscript
J Biomed Inform. Author manuscript; available in PMC 2018 April 12.

Published in final edited form as:
J Biomed Inform. 2015 October ; 57: 288–307. doi:10.1016/j.jbi.2015.08.005.

Hypothesis Generation Using Network Structures on Community Health Center Cancer-Screening Performance

Timothy Jay Carney, PhD, MPH, MBA[1,7,8,*], **Geoffrey P. Morgan, PhD Candidate**[2], **Josette Jones, RN, PhD**[1], **Anna M. McDaniel, PhD, RN, FAAN**[1,3], **Michael Weaver, PhD, RN, FAAN**[3], **Bryan Weiner, PhD**[7], and **David A. Haggstrom, MD, MAS**[4,5,6]

[1]Indiana University School of Informatics (IUPUI)

[2]Carnegie Mellon University, Computational Analysis of Social and Organizational Systems (CASOS)

[3]Indiana University School of Nursing

[4]VA HSR&D Center of Excellence on Implementing Evidence-based Practice Roudebush VA Medical Center

[5]Division of General Internal Medicine, Indiana University School of Medicine

[6]IU Center for Health Services and Outcomes Research, Regenstrief Institute, Inc

[7]University of North Carolina, Gillings School of Global Public Health

[8]Global Health Equity Intelligence Collaborative (GHEIC), LLC

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 some 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 diffuse to others as repeatable best practices. One reason is that facility-level changes typically occur in dynamic organizational environments that are complex, adaptive, and unpredictable. This study seeks to better understand the factors that help shape community health center facility-level cancer screening performance over time. This study applies a computational modeling approach that combines principles of health services research, health informatics, network theory, and systems science. Methods In order to investigate the role of knowledge acquisition, retention, and sharing within the setting of the community health center and the effect of this role on the relationship between clinical decision support capabilities and improvement in cancer screening rate improvement, we

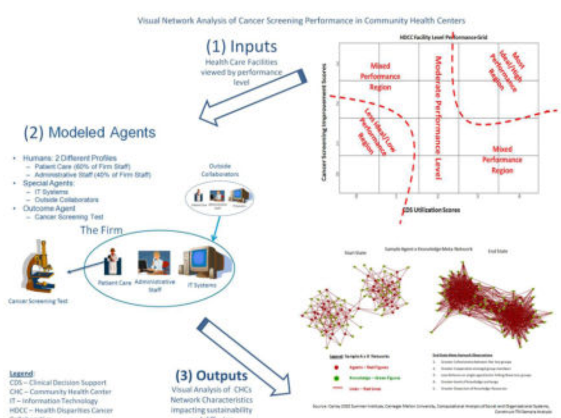
*Corresponding author: Timothy Jay Carney, tcarney@live.unc.edu.

STATEMENT OF CONFLICTS OF INTEREST

The authors declare that they have no conflicts of interests.

employed Construct TM to create simulated community health centers using previous collected point-in-time survey data. Construct TM is a multi-agent model of network evolution. Social, knowledge, and belief networks co-evolve. Groups and organizations are treated as complex systems, thus capturing the variability in human and organizational factors. In Construct TM, individuals and groups interact communicate, learn, and make decisions in a continuous cycle. Data from the survey was used to create high-performing simulated community health centers and low-performing ones based on extent of both computer decision support use and cancer-screening rates. Results Our 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 unconnected agents 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. Our study showed 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 and thereby enhance sustainability of long-term strategic improvement efforts. Our results revealed a strategic profile for community health center cancer screening improvement over a projected 10-year simulated period. The use of computational modeling and simulation allows for additional inferential knowledge to drawn from existing data examining organizational performance in increasingly complex environments.

Graphical abstract



Keywords

Computational Modeling; Simulation; Community Health Center; Systems-Thinking; Cancer Screening; Network Theory; Learning Health System

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 towards 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.[1915] 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 simulate the gain, spread, retention, and loss of the organization's knowledge over time through probabilistically determined informational exchanges, which are termed *interactions*.[20] This allows for evaluation of intermediary and long-term effects of an intervention on the knowledge component of an organization through virtual experiments. Entities within Construct-TM are known as *agents* (human or non-human) and interact due to (1) homophily, or similarity, between agents; (2) acquisition of new knowledge (i.e. learning); and (3) explicit information search (i.e. research).[20] In a simulation, an agent's knowledge can be inaccurate or incomplete, resulting in a less-than-optimal behavior and changed interaction probabilities with other agents in the network. Attributes assigned to or acquired by agents within a Construct-TM simulation are:[20]

- *Knowledge.* Defined by Construct-TM's user and represented mathematically as a binary or real value. It can be learnt or forgotten by agents at differing rates.
- *Beliefs.* An agreement with a principle area-function of current knowledge, prior beliefs, and simulation parameters, composition of the interaction sphere (a simulation parameter), and the influencing ability of others in conjunction with the susceptibility to influence knowledge.
- *Tasks.* Actions taken, possibly based on knowledge and beliefs.
- *Influence.* Interactions with other agents: i.e. initiating or receiving communication. Aside from homophily, the principle of influenceability is a prime driver of agent interactions within Construct-TM.
- *Socio-demographic Attributes.* Proximity measures affecting the likelihood of interaction. These include static factors (physical, socio-demographic, and/or social similarity weights) and knowledge factors (based on knowledge similarity and/or knowledge expertise). These determine the selection of a second agent to accept a fact "known" to or in possession of another agent.

2.2 Informing Construct-TM with Organizational Performance Survey Items

We used Construct-TM to create *virtual CHCs* similar to the virtual design team described by Jin and colleagues,[21–23] who used a "systematic" design of "organizational structures that relied upon abstracted descriptions of organizational tasks and activities" in a simulated model of the team.[22] The "virtual design team" pioneered the use of a simulated model to mimic the behavior of a full-scale, real-life organization.[22] We constructed our virtual CHC by relying on 37 summary measures derived from the 99 unique NCI/HRSA HDCC survey items. The HDCC survey provided details on cancer screening from the perspective of (1) organizational and/or practice setting factors, (2) provider factors shaping beliefs, attitudes, and behaviors, and (3) patient-population characteristics.

For an effective application of our method, the prerequisites include (1) availability of sufficient data on strategic planning to capture organizational performance and (2) a set of predefined metrics, which are available with CHCs and other healthcare organizations. We also assume this data is gathered largely to provide a retrospective view of activity for a specified time period, after which assumptions can be made about future events. However, in light of two primary issues specific to CHCs cancer screening facility-level performance stated above, sustainability and diffusion remain a problem and negatively impact long-term goals of reduction in cancer health disparities within vulnerable populations. We assume there exists hidden information in the strategic-planning data which is used inadequately to design future CHC practices. To address this issue, we designed a self-contained study to examine organizational behavior and the usefulness of the available data in extracting additional insights. An additional challenge involves the conversion of organization-performance assessments into usable network data using probabilistic-modeling algorithms within Construct-TM to extract data involving person-to-person contacts for interaction networks, person-to-resource/knowledge contacts for knowledge networks, etc.

Our study is self-contained, as the performance boundaries are limited strictly to the cohort of CHCs used. The terminology of "high-performance" and "low-performance" are not absolute and are used relative to each other (based upon their HDCC survey responses). Additionally, we utilize four distinct states for comparison: (1) low-performing CHC-beginning state, (2) high-performing CHC-beginning state, (3) low-performing CHC-end state, and (4) high-performing CHC-end state. The beginning states are derived from the existing data, and the end states are based on probabilistic modeling of interaction, knowledge utilization, learning, and the ability to influence others or to be influenced. The high-performing CHC-end state represents the ideal state that any of the CHCs can assume within the given timeframe of ten years, while low-performing CHC-beginning state represents the least-desirable state. Two major challenges to achieving *virtual CHCs* include the following: (1) if successful simulation of performance level behavior throughout the ten-year simulated period can be achieved to reflect CHCs using the existing HDCC survey variables, and (2) if observable differences detected within the network structures can be hypothesized as facilitating/inhibiting sustainability and diffusion.

The first challenge involves the assignment of a correct set of variables to define the behaviors and practices of each agent specifically and the performance level generally (reviewed below in Methods). The second challenge pertains to determining the extent to which the experiment could yield visual distinctions in relation to the performance levels and add value to traditional assessments of CHCs cancer-screening performance (HDCC survey, in our case). A successful experiment would yield critical intelligence on the specific data set and variables gathered within the HDCC cancer-screening performance survey, along with the parametric boundaries that best describe and predict the facility-level and agent behavior over time. A successful model will enable the identification of structural differences linked to performance levels, aiding in the development of strategies for the diffusion of best practices and innovation. We argue that the current HDCC-survey data cannot model successfully the future performance of CHCs with a traditional statistical model alone.

3. Evaluating Performance via Network Visualization

Network analysis is used typically to examine interactions among and between entities (referred as agents or actors) within a predefined boundary or space. In our modeling, an agent (A) can be human or non-human models of interaction between two or more people, between a person and a knowledge resource (K), or between core knowledge resources and critical tasks (T) that can be created.

In the present study, we limit our investigation to the knowledge-sharing network to evaluate knowledge-sharing practices between agents (A × A) and the knowledge-resource network (A × K), as well as to observe access to knowledge resources among agents in the network. We rely largely on visual analysis of network diagrams to highlight the differences between performance levels. The basic construction of a network is characterized by nodes (represented by an individual agent or resource) and linkages (the connecting lines among agents or between agents and resources). Network relations among agents and resources are illustrated by sample network diagrams recreated from data provided by the Center for Computational Analysis of Social and Organizational Systems (CASOS) at Carnegie Mellon University.[18] Figures 1a–1d illustrate the initial- and final-state sample network diagrams and help to contextualize our use of network measures.

We limit our analysis to six distinct network measures: density, cliques, clusters, connectedness, cohesion and symmetry. Our networks are bimodal, with agents connected to knowledge. Therefore, concepts such as *betweenness centrality* and *eigenvector centrality* are not appropriate entirely for describing these networks, although they are useful analogical concepts.

3.1 Network Density

Network density is defined as the proportion of ties actually present from the total number of possible ties. When depicted using a force-directed layout, highly dense networks are presented as tightly packed, and the node-level measure of this quantity is called degree.

3.2 Cliques and Clusters

A clique represents a concentrated or localized density of network nodes, represented as agents or resources. The formal definition of a clique is a group involving "the maximum number of actors who have all possible ties present among themselves." The smallest cliques can consist of as few as two nodes (referred as dyads) or can grow to form much larger, closely connected groups within the network. Cliques provide insight regarding access to key resources or actors within the network, the level of proximity or isolation between individuals or resources to others within the same network, and the level of overlap of these agents/actors within the network. Clique analysis is critical when examining diffusion and adoption of innovation studies. Innovations often have difficulty penetrating a clique initially, but then spread rapidly.[24,25]

3.3 Connectedness and Cohesion

While network density can be observed by examining the arrangement of nodes and their relative distance to each other in the network, connectedness and cohesion are used to examine the arrangement of linkages within a network. Connectedness measures how many linkages one agent has to other agents within the network, both in-group and out-group. This is examined visually by observing the number of linkages in the network. Here we assume that more visible linkages refer to a network displaying increased connectedness.

Within our sample network, connectedness can be observed in terms of *within*-cluster connectedness and *between*-cluster connectedness. As an example, we consider each cluster as representative of a distinct agent classification (e.g. general practitioners and specialists), and we examine the level of coordination among primary-care and specialty-care providers in sharing knowledge resources with each other and the patient in delivering care. In our sample network, we observe that in the initial state within each cluster, several sub-clusters or cliques display more connectedness, which reveals a less uniform arrangement of agents in the network. Also, the relative connectedness of agents to knowledge resources is distributed unevenly, making certain knowledge resources more accessible to some agents than to others. Analysis between clusters indicates that two core clusters in the initial state are connected only by a single agent, referred to as a *bridging agent*. A bridging agent links two or more distinct clusters; although this resembles the ideas of *betweenness*, these are multi-modal networks. A bridging agent has a unique collection of knowledge instead of social relationships.

Considering the example of primary care vs. specialists, the bridging agent might be a patient navigator or care coordinator. Elimination of a bridging agent compromises the integrity of interaction across the clusters. In the end state, connectedness is increased greatly over the start state and with uniform access to critical knowledge resources. In our network analysis, more connectedness is considered favorable.

Cohesion defines not only the number of linkages, but also the quality of the linkages (defined as agent to agent sharing, cooperation, and collaboration). Cohesion is, in part, observed in patterns of clusters of cliques formed. Clusters and cliques must be examined within the context of phenomena in question.

Consistent with our analysis, the end state in our sample network suggests a heightened state of all of these factors, evident from the strong centralized cluster or clique encompassing nearly all agents, signifying strengthened ties amongst agents and a strong, tightly knit environment. Additionally, the knowledge resources are more accessible, and the two larger clusters within the end state have more ties between them and no longer rely upon a single bridging agent. The existence of knowledge nodes between the two clusters indicates shared knowledge resources that are accessible to both clusters. In the start state, the knowledge resources were only available within each cluster, and the bridging agent served as a curator for inter-cluster knowledge transfer. Within our analysis we assume that greater patterns of cohesion indicate a more favorable state than lesser cohesion.

3.4 Network Symmetry

There is an ongoing debate whether a symmetrical network represents a more favorable state than a less symmetrical one. Within the context of sustainability and diffusion, we assume that a symmetrical network allows for enhanced information flow, improved knowledge utilization and sharing, and increased interactions among agents. Network symmetry can be observed visually and serves to combine several concepts such as the proximity of nodes, the arrangement of the connection of nodes, and the pathways agents/resources must travel to reach targeted agents/resources. While analyzing for symmetry, a visual inspection is performed for areas of the network exhibiting vulnerability in terms of the need for or absence of potential bridging agents, randomness in configuration vs. highly predictable/ repeatable patterns, and relative distribution of agents and resources throughout the network. The relative distribution of nodes in a network is critical to understand the overall health of a network. Ergo, the principle that the network is only as strong as its weakest link (or node) is an appropriate reference.

In our sample network (Fig 1c), the start state displays knowledge resources (green) and agents (red) randomly scattered and without any discernible patterns. While we observe two large clusters and several observable cliques, the overall start-state network represents a highly asymmetrical arrangement of agents and knowledge resources, resulting in an uneven accessibility of knowledge resources to the neighboring cliques and no accessibility to other members of the network. The same applies to sharing between agents, as some agents might have access to neighboring agents but limited or no access to other agents across the network. A healthy organizational network allows a consistent distribution of and access to knowledge resources and agents across the network, assuming that collaboration, cooperation, and cohesion are desired to interact either directly or through connections in the network.

We also observe a “crowning” arrangement in the end state produced by the knowledge resources, indicating that the agents are surrounded amply by knowledge resources to support the activity of the centralized cluster of agents. The centralized clustering of agents and the crowning effect of the knowledge resources represent an easily identifiable pattern, which is predictable and symmetrical. The challenges of sustainability and diffusion throughout the network are likely to be achieved in a network displaying end-state configurations instead of start-state configurations.

By employing the sample network diagrams, we illustrate that visual analysis of network diagrams can assess network density, cliques and clusters, cohesion and connectedness, and symmetry. A comparison of the start- and end-state sample networks can detect an evolution in network cohesion, cooperation, collaboration among agents; decreased vulnerability (less reliance on a single bridging agent); increased access and exchange of knowledge; and more evenly dispersed patterns of knowledge-resource availability with the end state representation as highly favorable.

A similar methodology evaluates CHCs cancer-screening performance and generates hypotheses on factors shaping the long-term sustainability of quality-improvement efforts (e.g. HDCC) and the increased diffusion of best practices and innovations for the improvement of cancer-screening efforts.

4. Methods: Modeling Community Health Center Behavior with Construct-TM

The primary aim of this computational analysis is to generate hypotheses by examining the outcomes of a cohort of *virtual CHCs*. Outcomes are examined through visual analysis of the resulting interactions and knowledge networks of CHCs. We conducted a facility-level comparison, expressed as high vs. low performance levels for activities involved in the screening of breast, cervical, and colorectal cancer. Our “between” analysis focuses on comparing the two performance levels to determine observable structural differences that account for a lack of sustainability of quality-improvement efforts over time and for the slow uptake and diffusion of innovation and best practices throughout the network of CHCs.

To achieve these objectives, we targeted the following three tasks: (1) to convert a point-in-time survey of organizational, provider, and patient characteristics associated with cancer-screening performance into a data source for dynamic network-simulation modeling, (2) to develop a sufficiently sensitive simulation model to discriminate between performance levels based upon the 37 summary measures, and (3) to project these simulated behaviors over a ten-year time period.

4.1 Data Preparation for Entry

As published previously,[26] CHC data collected during the HDCC survey was mapped with 37 summary measures concerning organizational and/or practice-setting factors, provider characteristics, and patient characteristics to explain CHCs cancer-screening behaviors.[26] For the current study, these summary measures were mapped with the following Construct-TM categories to yield a set of formal definitions and parameters governing agent behavior throughout the simulation: knowledge, task, agent, and belief.[15,19]

4.2 Study Sample

Data was retrieved from the NCI/HRSA HDCC 2006 survey comprised of a representative sampling of 44 CHCs. Of these, 22 CHCs were identified as participants within the NCI/HRSA sponsored 2003–2005 cancer-screening improvement collaboration, while the other 22 CHCs did not participate. The HDCC survey measured the impact of HDCC participation on CHCs overall cancer-screening improvement over a 12-month period. The sample was biased intentionally toward high-performers in both participants and non-participants to ensure an observable effect was detected if it existed. A secondary analysis was performed to determine the extent of impact of the clinical decision support (CDS) as an innovative best practice (with the use of clinical reminders, use of provider prompts at point-of-care, and generation of automated patient results for providers) on facility-level cancer-screening improvement scores.[27] Although our previous results indicate a positive correlation between HDCC participation and CDS use, traditional statistical analysis could not find any correlation between the use of CDS and its impact on cancer-screening improvement in

CHCs. We were also unable to observe continued rates of improvement in cancer screening or sustained uptake of innovative technologies.[27]

Therefore, we created a composite outcome measure of cancer screening improvement and CDS utilization to represent the overall performance of CHCs, which was later used to instantiate our *virtual CHCs*. These virtual CHCs served as a basis for computational analysis. For the computational model, each CHC was assigned a composite performance level based on its graph location (Fig. 2). The CDS measure for the intensity-of-use (ranging from 0 to 4) formed the graph's x-axis, and the cancer-screening improvement score (ranging from 0 to 3) formed its y-axis. The resulting matrix was divided into *high*, *medium*, and *low* regions for each measure. Based on its position in the graph, a CHC could be assigned a qualitative, two-part coordinate, which formed the unit of analysis of virtual CHCs within Construct-TM. Figures 2 and 3 represent this distribution and the plot matrix of CHCs, respectively. Our current analysis (as well as the previous computational-analysis study) does not discriminate on the basis of HDCC participation.[27] Our initial statistical analysis of CDS uptake tested the effects of both HDCC cumulative exposure and current membership, influencing uptake.[26] The results indicate that cumulative exposure to collaborative activities had a higher impact on outcomes when compared to the current collaborative membership status.[26]

4.3 Identifying Agents Within the Virtual CHCs

Initially, to adapt the survey data to Construct-TM, five entities (either positions or functional units within a CHC) were selected to act as agents in the simulation. Five survey-respondent groupings identified in the original survey served as a basis for the simulated agents: (1) director (CEO), (2) chief financial officer (CFO), (3) general staff, (4) provider, and (5) chief information officer (CIO). Since several survey items addressed the utilization of outside agreements with medical specialists via contracting, collaborative agreements, and sharing of best practices among CHCs, an agent classification of *outside collaborator* was added (ignoring the membership status at the time of the survey). As the study involves cancer screening, an agent classification for the cancer-screening test itself was also included. The agent set thus consisted of (1) firm view–administrative, (2) firm view–clinical care, (3) outside collaborators, (4) IT systems–CDS, and (5) cancer-screening tests (CSTs).

4.3.1 Agent tasks, knowledge, and beliefs—After initiating the agents in the virtual CHC's, agents were characterized according to the tasks they performed, knowledge they possessed/shared, and their beliefs. This was accomplished using the data set from the HDCC survey (37 summary measures) according to the following definitions:

- *Task* assignment: an action performed by the agent.
- *Knowledge* assignment: information in possession of an agent.
- *Belief* assignment: principles or assertions believed by an agent to be true or false.

The logic of question(s) composing one of the 37 summary measures determined its category assignment and were validated by an internal advisory committee of subject-matter experts (SMEs). Subsets of the 37 summary measures were used to define agent behavioral characteristics and performance-level practices in the simulation. Appendix I enlists the variable subsets used to define each agent classification in our analysis. Key assumptions and factors shaping agent performance of tasks related to cancer screening, as well as opportunities identified for an exchange of ideas, learning, and shaping of beliefs (referred to as homophily or similarity knowledge), are also enlisted.

5. Construct-TM Glossary and Definitions

The Construct-TM Variable Glossary highlighted in Appendix II outlines the full list of 37 summary measures used to define the simulation, along with their Construct-TM coded name.

5.1 Agent Definitions

For each of the five agents described above (Administrative, Patient Care, IT Systems, Outside Collaborators, and CSTs), useful definitions were created based on a set of critical assumptions. These include the following:

- One IT system actor was sufficient to represent all available technology capabilities for CDS used in support of cancer-screening activities.
- One outside collaborator agent was sufficient to represent all official outside collaborations.
- Three types of CSTs were in existence: (1) colorectal, (2) cervical, and (3) breast cancer.
- The total number of people equals 100× the normalized value for the financial budget (as a mathematical relationship between organizational size and budget).
- Of the personnel in CHCs, 60% contributed to the agent *Patient Care*, and 40% contributed to agent *Administrative Staff* (these arbitrary percentages were based on the proportion of survey responders labeled as Administrative vs. Clinical and are not an actual representation of CHCs personnel distribution).

The definition of each agent comprised the logic used to determine the start value (current knowledge, connections, and interactions) for the agent and for his/her end value (post-simulation knowledge, learning, and connections) within the simulated range of activities referred to as an array. The agents were set up to allow Construct-TM to read the value for each agent in the array and to move automatically to the next agent classification for input into the simulation. The start of one agent value is defined where the previous agent ends, as seen in Table 1.

5.2 Knowledge Definitions

Knowledge definitions were based on the assumption that all tasks are roughly equivalent in complexity (i.e. each task has the same number of “bits” in Construct-TM, informing it throughout the simulation). Bits establish the level of expertise or knowledge saturation a particular agent may have (ranging from 0% to 100%). Within the same simulation, a complex task has more bits informing its execution.[20] For this Construct-TM virtual experiment, we assumed each task was roughly equivalent in complexity and therefore had the same knowledge bits (50) assigned to each. This equivalence was chosen due to a lack of information in discriminating the tasks (i.e. the relative importance of the variables within the HDCC survey items among respondents was not known). As with other simulation assumptions, if more information was available, the assumption could be revised. For now, the definitions for knowledge reveal the start and end of the 50-bit index assigned to each knowledge element as task knowledge and homophily knowledge (see Appendix I). Table 2 lists the system-wide schema for all knowledge that the agents must possess or sources from which it can be obtained in conducting cancer-screening practices.

5.3 Task Definitions

The task definitions associated with each agent's knowledge consisted of assigning an index number for Construct-TM to read it into the array and to provide a start and end point for each task in the simulation. Appendix III enlists a complete list of tasks used in the present analysis and defines the tasks that agents perform or factors that might influence the performance in conducting cancer screening. Agents are not expected to be able to perform every task of the organization.

5.4 Simulation of Performance Measures

The next step involved defining the simulation's outcome measures and determining the role of Construct-TM's output in generating hypotheses. In this analysis, simulation outcomes focused on organizational-environmental learning as a function of either the task performance or the absorption of knowledge, both of which were measured in terms of the facility-level improvements in cancer screening, defined as the chief outcome of interest. The Construct-TM model's primary performance measures were knowledge gains and task capabilities, as evidenced by interactions and knowledge-absorption rates. Although CDS application was also critical, we viewed it as a driver of overall cancer-screening performance and as an artifact of diffusion.

5.5 Data Preparation

To convert HDCC summary measures into useful inputs for Construct-TM, inputs were provided in the form of Extensible Markup Language (XML). We developed an XML code generator in Excel, and this Excel spreadsheet served as a functional, special-purpose graphical user interface (GUI) for Construct-TM. A Construct-TM input deck comprised of four critical elements: *variables* (factors being studied or system components), *parameters* (the mathematical/system boundaries that the variables can assume), *nodes* (the representation variables within the network/elements of the system), and *networks* (the overall representation of phenomena in the studied context or system dynamics). This Excel XML generator was used to create a valid XML code for each of these elements, allowing us to focus on interpreting HDCC measures rather than considering the basic syntax of XML.

5.6 Visualizing Construct-TM Output

The CASOS tool-set contains a network-visualizer tool called ORA® (originally Organizational Risk Analyzer, but now used as a psuedo-acronym)[28], which permits visual and qualitative analysis of networks with over 100 network-analysis measures (e.g. connectedness, *between-ness*, density, centrality, etc.). We employed a small subset of these measures, which were selected because of their relevance to our hypothesis-generation exercise. Construct-TM output was inserted into ORA® to generate network diagrams for visual inspection of networks to observe changes in the visual displays of the networks over the 10-year simulation period. Patterns of interest included knowledge/resource utilization within CHCs in the agent-by-knowledge (A × K) network and those in the agent-by-agent (A × A) network, representing knowledge-sharing patterns among network agents within CHCs. Model Year-1 represented the beginning of the simulation period (beginning state), while Year-10 represented the end state for each performance level with respect to network measures (density, clusters and cliques, cohesion and connectedness, and symmetry). In this paper, however, we present only the comparisons of the two most extreme conditions (performance levels): high/high (HH) and low/low (LL). As mentioned previously, the composite measure was scored for cancer-screening improvement (high, medium, or low) plus CDS use (high, medium, or low).

6. Results

6.1 Community Health Center Characteristics

HDCC Sample Means and Standard Deviations for Summary Measures by Performance Level

6.2 Diffusion and Task Capability over Time

Because we are using a simulation model, we can examine outputs at each time-period of the simulation. Figure 4 demonstrates how HH and LL groups interact and evolve differently over the course of the simulation.

The diffusion curves show a steady increase in knowledge in each group, but the task capability patterns, defined as the number of knowledge-groups for which an agent has more than half of the knowledge available for the task, are non-linear. Despite being related to knowledge diffusion, task capability is sensitive to different interaction patterns.

In general, LL firms tend to aggregate task capability regularly over time, indicating that agents learn information about a variety of tasks over the course of time, implying that the actors in LL firms seek to be generalists. The HH firms, on the other hand, show different rates of growth over the course of the simulation, suggesting that HH actors tend to interact within their groups and that new knowledge acquired by an actor is spread rapidly through their membership group. Thus, HH actors tend to be more specialized.

6.3 Outcomes at Year 10: The Knowledge/Resource Utilization Network (A × K)

Figures 5a and 5b represent the network diagrams produced in ORA®, highlighting our agent-by-knowledge (A × K) network at simulation start time (Year-1, Figure 5a) and at simulation end time (Year-10, Figure 5b). The agent classes for the simulation are shown in red, while knowledge elements are in green. Connections between agent classes and knowledge elements are represented by blue lines.

Both HH and LL Year-1 networks reveal four distinct agent clusters or cliques, characteristic of interaction among similar agents. Uneven distribution of knowledge resources and unconnected agents and knowledge resources are evident. Unconnected agents within a network diagram represent a lack of connection to the core group/key resources, while unconnected knowledge elements represent unused or underused knowledge resources (issues of access or being outdated), with each unconnected agent and knowledge element contributing to a less-than-ideal state.

The Year-10 networks exhibit an evolution of HH and LL networks toward ideal states, reflected by higher cohesion between the agent (represented by the central red cluster) and the knowledge-resource (represented by the green clusters around agents). HH displays only one major agent cluster in the network center, suggesting greater levels of overall agent cohesion, while LL has the presence of a secondary agent cluster, a potential clique. Interestingly, although Year-10 LL network has a higher overall density than Year-10's HH network, Year-10 LL network configuration is less than ideal because of a wide distribution of knowledge elements.

6.4 Outcomes at Year 10: The Knowledge-Sharing Network (A × A)

Figures 6a and 6b represent the simulated evolution of knowledge-sharing practices within the HH and LL networks over the 10-year period. We measured the degree of communication between network agents or the sharing of core knowledge resources essential to the performance of the cancer-screening task.

The agent classes defined for A × A network are the same as in the A × K network (in Figures 5a and 5b). However, the A × A knowledge sharing network was used to infer homophily relations between actors (colored by their membership group). As in the A × K network, greater density is considered favorable with respect to information dissemination, sharing, and exchange, unless such increased density is coupled with structural network characteristics that inhibit sharing of core knowledge resources.

Both HH and LL networks show nearly identical patterns of knowledge sharing in Year-1, with three distinct agent clusters: patient-care (clinical) staff (green), administrative staff (blue), and the cancer-screening-test agents (red). Administrative agents serve as a knowledge bridge between patient-care agents and cancer-screening tests. LL firms also appear larger, and the Year-1 LL network displays greater density than HH, consistent with previous results.

When simulation is initiated, there is little difference between the two networks. However, the simulation results for Year-10 display significant differences in network configurations. Both firms develop a single large cluster, as information is gained by all parties. However, relative specialization of HH firms is evident in the cohesive patient-care and administrative blocks, while LL firms have a single core-cluster of firm members where patient-care and admin agents are dispersed randomly.

7. Discussion

The reported analysis provides possible alternative hypotheses explaining facility-level cancer-screening performance to examine in future work. To explore the sustainability of quality-improvement efforts and the diffusion of innovative/best practices, we use a well-regarded simulation engine to instantiate virtual CHCs and then to examine the network characteristics of the CHCs. We limited our network-analysis measures to network density, cohesion, presence of cliques, the level of connectedness, and symmetry, as previous work by Bruque and colleagues suggests that network size, network density, and the strength of information ties serve as predictors of the ability to adapt and that a dense information network displays increased adaption as long as network members use information to resolve doubts, solicit opinions, and deepen understanding of the new system or of existing strategies for improvement.[29] Such a system displays greater sharing, a more effective use of knowledge resources, and a greater capacity for information-exchange, growth, and evolution.

This study employed density as the single objective measure but also included the subjective measures like network cohesion, cliques, and collaboration to provide additional perspective. In addition, network symmetry and overall accessibility to knowledge (in relation to other members of the network) were also considered.

We show that the visual comparison of high- and low-ranked CHCs network configurations for Year-10 HH firms possess an ideal agent-to-knowledge-resource configuration in comparison to LL, despite lower levels of overall network density. The LL end-state configuration did not facilitate streamlined and efficient knowledge/resource utilization nor encourage greater degrees of knowledge-sharing outside of the centralized agent cluster, confirming that network density alone was not an ideal indicator of organizational performance.

These findings were consistent with previous studies for all parameters except network density. This is not surprising, as there are numerous configurations possible for a network to assume that may display relatively identical density, with some more effective than others in the transfer of knowledge.

Overall, the higher ranked CHCs demonstrate greater network cohesion with streamlined and efficient forms of collaboration (expressed as a function of the knowledge-sharing network's symmetry) and with fewer separate group clustering or cliques. For connectedness, results were mixed; higher-ranked CHCs in Year-10 have fewer unconnected knowledge resources than lower-ranked centers, but they also have greater number of unconnected agents than the Year-10 lower-ranked centers, indicating that performance might have a better correlation with knowledge/resource utilization than with agent connectedness.[29]

We demonstrate that network elements indicate performance level over time and provide insights into organizational ability to learn, exchange information, and adapt over time. Network configurations with respect to CHCs' performance allow for hypotheses to be drawn and tested in future experiments to examine sustainability and diffusion. We assert that long-term sustainability and diffusion may be positively correlated with network characteristics such cohesion and connectedness, cliques and clusters, and symmetry and that they are correlated less tightly with network density. We also generated hypotheses regarding the coupling effect of command and control structures and agent-to-agent knowledge-sharing practices as predictors of long-term sustainability and diffusion.[25] This present study suggests that the mere presence of a highly organized quality-improvement effort may indeed yield positive results in the short-term, but progress sustainability and diffusion of practices within and across agencies are shaped by factors with fewer ties to processes and outcomes.

Visual inspection of the network diagrams strongly suggests that the network-evolution simulation model presented here can provide significant insight into the organizational performance of future CHCs and can provide a basis for continued hypothesis generation and testing on ways to sustain effort and to encourage best practices. Finally, through computational modeling, we could extract additional value and insights from the data set beyond the traditional analysis of a single point-in-time survey.

7.1 Simulation-Model Validation

Sargent describes the process through which system theories about the world are incorporated into simulations to fulfill system-level experimental objectives and to produce results for hypothesis generation.[30] Although real-world inferences are a natural progression in any computational-modeling exercise, model validation is a necessary prelude.

We addressed issues of learning, adaptation, and evolution of social and organizational systems by examining the problem of facility-level performance with respect to both clinical-decision support and cancer-screening rates. We used both a reductionist statistical model based on real-world data and a simulation. The model validation categories are (1) internal, (2) parameter, (3) process, (4) face, (5) pattern, (6) content, (7) external, and (8) theoretical.

Internal validity determines if the computer code is correct and error-free.[31] It employs strategies to ensure that all steps including data collection, data entry (via CASOS's Excel-based XML code generator worksheet), and data transformations in the study maintained a high degree of accuracy. Furthermore, each series of statements was tested and debugged iteratively.

Parameter validity involves correct assignment of the simulation parameters to ensure that each of the 37 summary measures derived were mapped properly to the appropriate simulation category (e.g. task, belief, measure of knowledge, agent). As available data did not specify explicit interaction partners, we used statistical means to determine and to define the average probability of an agent's knowledge allocation (represented as knowledge bits, k) and relied on Construct's homophily and expertise-seeking drives to suggest interactions.

This study posited three representations of the “Cancer-Screening Test” agent (e.g. colorectal, breast, and cervical) for each clinic, all of which represented the clinic’s competency in the key tests of interest. We measured the saturation of knowledge in these three agents over time, allowing for a direct comparison across all test cases.

Process validity determines if the study is conducted in a dependable, competent manner and if efforts are not focused on appraising existing practices.[32] It is understood as the extent to which actions and thought processes of test takers or survey responders demonstrated that they understood the construct in the same way the researchers intended.[33] The HDCC survey performed by Haggstrom and colleagues generated data used in this simulation study.[3] Therefore, SME assessment was employed to validate the following aspects: (1) the original researcher’s interpretation of summary measure data as agent, task, belief, or knowledge measure, (2) the assignment of specific summary measures to describe the behavior of each of the five agents used in the simulation, and (3) the assignment of each of the 44 CHCs to one of five performance levels used in the simulation. The experts also ensured that the logic of survey questions and/or intent of the primary data collector were maintained throughout the development of the simulation experiment. Process validity can be extended to the issue of content validity, defined by Merrill and colleagues as the principle guide for formulating survey questions’ specific relationships in a predefined network.[34,35] Therefore, the HDCC survey data used in this study can be considered a valid representation of the agents, knowledge, tasks, and beliefs of CHC workers.

Regarding pattern validity and face validity, study results for the respective performance levels are relative to their initial states. *Pattern validity*, also called relational equivalence,[36] is defined as the degree to which patterns in the data reflect observed results. Closely related is *face validity*, which is essential for the study to be considered a reasonable representation of reality.[37] In the methods section we describe how two measures—facility-level CDS utilization scores and cancer-screening improvement-rate scores—were used to construct a composite representation of CHC performance and how each facility was designated to one of the five performance levels based on this composite measure. We initiated our study based on the assumption that better-performing facilities exhibit higher patterns of learning and knowledge absorption over time, and the simulation network visualizations confirmed these expectations, providing greater insight into the impact of knowledge sharing, resource utilization, group cohesion, and network connections on learning over time. In all instances, results obtained were reasonable representations of CHC performance, as reflected by the HDCC survey.

Merrill and colleagues raised the issues of *external validity*, or generalizability of study results, with respect to network analysis: “External validity refers to the adequacy and accuracy of the computational model in matching real world data.” [31] Merrill and colleagues suggest that the model can be validated by correlating network findings with observed data.[34,38] Although this does not imply a generalization of the model for real-world implementation, it does imply that the model—if evaluated in the context of stakeholders knowledgeable in both CDS and cancer-screening-related organizational structure and operations—may enhance overall facility-level learning and knowledge absorption over time as related to the two outcomes.

Finally, the issue of *theoretical* validity represents the ability of the findings to reflect current theory (i.e. that model assumptions fit the problem and that model instances were specified appropriately). Sargent describes how model confidence is a function of the cost of conducting the test and the value of the model to a predefined user, stating that "The cost of model validation is usually quite significant, especially when extremely high model confidence is required."[30]

We chose specific tests for simulation modeling based upon the following criteria: (1) hypothesized relationships and statistical inferences drawn from the statistical model (the relatedness of the antecedents to both the proximal and distal outcomes),[26] (2) recognizable patterns in the data set of CHC characteristics (e.g. the grouping of performance levels), and (3) expert guidance from SMEs (e.g. which antecedents should be used to inform which agents).

While no formal cost curve was developed for this study, cost-benefit can be understood in terms of the value added to data obtained from a single point-in-time source and of using this data for future, long term projections. The simulated model of the 2006 HDCC survey added value to the survey data, projecting the 10-year outcomes (assuming that the real-world interactions observed in the point-in-time data remained constant over this time horizon). This highly specific set of chosen tests examined the following: (1) rate of learning or knowledge absorption, and (2) patterns of cohesion, interactions, and interconnections expressed as network diagrams. These topics of interests added value to the long-term strategic-planning efforts aimed at addressing CHC performance objectives for cancer screening and CDS.

8. Conclusions

This study demonstrated that a healthcare facility, defined by 37 summary measures obtained from a previous organizational survey of cancer-screening, can be described as a learning organization and a function of the following: (1) knowledge/resource utilization by key agents, and (2) agent-to-agent sharing of core knowledge to support cancer-screening quality and improvement efforts.[29,39–41]

This research successfully produced a meaningful virtual experiment in computational-modeling, mimicking reported activity within the CHC sample associated with the cancer-screening test. We demonstrate that our model is (1) sensitive enough to differentiate between high- and low-performing CHCs, (2) could identify observable structural differences between high and low performers, and (3) can establish a research platform to explore hypothesis generation and testing to achieve sustained improvement and diffusion of best practices after traditional quality-improvement efforts are completed. Our model views the healthcare organization as a complex adaptive entity,[40–43] reinforcing the findings of previous studies associating high-performing firms with greater learning, intuition, or knowledge absorption in clinical knowledge-management practices. Our research effort was designed to generate a series of hypotheses that can be developed and tested in future experiments to contribute to more informed interventions.

9. Limitations

We were unable to apply every combination of summary measures to each agent class within the simulation, limiting the generalization of the model and its applicability. Additionally, the rigorous selection of the summary measure describing agent behavior dramatically limited the number of ways the agent could learn, interact, and evolve within the simulation. Future research might explore less-rigid criteria for inclusion of variables to account for unpredictable changes in the data, employ a more-sophisticated algorithm capable of testing any or all combinations of variables, and allow for agents in the simulation to forget, a capability not allowed in this simulation. We examined knowledge acquisition alone and assumed that, once acquired, the knowledge was retained through the remainder of the simulation.

Acknowledgments

This research was supported by Indiana University (IUPUI) School of Nursing Training in Research for Behavioral Oncology and Cancer Control Program (TRBOCC) under The National Cancer Institute Pre-Doctoral Fellowship R25-CA117865-04.

The manuscript is supported by The University of North Carolina Gillings School of Public Health and The Lineberger Comprehensive Cancer Center - Cancer Health Disparities Training Fellowship T32CA128582 and The Carolina Community Network Center to Reduce Cancer Health Disparities Diversity Supplement 3U54CA153602.

Dr. Stephen H. Taplin and Dr. Stephen B. Clauser allowed access to the original data set, providing the foundation for this research through their previous findings, and helped in developing a future direction of this research.

References

1. Taplin SH, Haggstrom D, Jacobs T, et al. Implementing Colorectal Cancer Screening in Community Health Centers: Addressing Cancer Health Disparities Through a Regional Cancer Collaborative. Medical care. 2008; 46:S74–S83. DOI: 10.1097/MLR.0b013e31817fdf68 [PubMed: 18725837]
2. Finney Rutten LJ, Nelson DE, Meissner HI. Examination of population-wide trends in barriers to cancer screening from a diffusion of innovation perspective (1987–2000). Preventive medicine. 2004; 38:258. [PubMed: 14766107]
3. Haggstrom DA, Taplin SH, Monahan P, Clauser S. Chronic Care Model implementation for cancer screening and follow-up in community health centers. Journal of health care for the poor and underserved. 2012; 23:49–66.
4. [Accessed March 2010] Cancer Screening Overview. 2009. at http://www.cancer.gov/cancertopics/pdq/screening/overview/HealthProfessional
5. Kilo CM. A framework for collaborative improvement: lessons from the Institute for Healthcare Improvement's Breakthrough Series. Quality management in health care. 1998; 6:1–13.
6. Asch SM, Baker DW, Keesey JW, et al. Does the collaborative model improve care for chronic heart failure? Medical care. 2005; 43:667– 75. [PubMed: 15970781]
7. Chin MH, Cook S, Drum ML, et al. Improving diabetes care in midwest community health centers with the health disparities collaborative. Diabetes Care. 2004; 27:2– 8. [PubMed: 14693957]
8. Kilo CM. A framework for collaborative improvement: lessons from the Institute for Healthcare Improvement's Breakthrough Series. Qual Manag Health Care. 1998; 6:1– 13.
9. Mittman BS. Creating the evidence base for quality improvement collaboratives. Annals of internal medicine. 2004; 140:897–901. [PubMed: 15172904]
10. Construct - A Multi-Agent Network Model for the Co-Evolution of Agents and Socio-Cultural Environments. Defense Technical Information Center; 2004. Accessed at http://handle.dtic.mil/100.2/ADA460028

11. Krackhardt, D., Carley, KM. PCANS model of structure in organizations. Pittsburgh, Pa: Carnegie Mellon University, Institute for Complex Engineered Systems; 1998.
12. Prietula, MJ., Carley, KM., Gasser, LG. Simulating organizations : computational models of institutions and groups. Menlo Park, CA: AAAI Press/MIT Press; 1998.
13. Burton, RM., Obel, B. Strategic organizational diagnosis and design : the dynamics of fit. Boston: Kluwer; 2004.
14. Kling R. Organizational Analysis in Computer Science. INFORMATION SOCIETY. 1993; 9:71.
15. Effken JA, Brewer BB, Patil A, Lamb GS, Verran JA, Carley KM. Using computational modeling to transform nursing data into actionable information. Journal of biomedical informatics. 2003; 36:351–61. [PubMed: 14643731]
16. Ilgen, DR., Hulin, CL. Computational modeling of behavior in organizations : the third scientific discipline. Washington, D.C: American Psychological Association; 2000.
17. Carley KM. On Generating Hypotheses Using Computer Simulations. SYSTEMS ENGINEERING -NEW YORK-. 1999; 2:69–77.
18. Center for Computational Analysis of Social and Organizational Systems. Institute for Software Research in the School of Computer Science at Carnegie Mellon University. 2009. (Accessed at http://www.casos.cs.cmu.edu/.)
19. Effken JA, Brewer BB, Patil A, Lamb GS, Verran JA, Carley K. Using OrgAhead, a computational modeling program, to improve patient care unit safety and quality outcomes. International journal of medical informatics. 2005; 74:7–8.
20. Specifying Agents in Construct. Defense Technical Information Center; 2009. (Accessed at http://handle.dtic.mil/100.2/ADA500804.)
21. Jin Y, Levitt RE, Christianson TR, Kunz JC. The Virtual Design Team: A Computer Simulation Framework for Studying Organizational Aspects of Concurrent Design. Simulation. 1995; 64:160.
22. Kunz JC, Christiansen TR, Cohen GP, Jin Y, Levitt RE. Articles - The Virtual Design Team. Communications of the ACM. 1998; 41:84.
23. Levitt RE, Thomsen J, Christiansen TR, Kunz JC, Jin Y, Nass C. Simulating Project Work Processes and Organizations: Toward a Micro-Contingency Theory of Organizational Design. Management Science. 1999; 45:1479–95.
24. Young HP. The dynamics of social innovation. Proceedings of the National Academy of Sciences of the United States of America. 2011; 108:21285–91. [PubMed: 22198762]
25. West E, Barron DN, Dowsett J, Newton JN. Hierarchies and cliques in the social networks of health care professionals: implications for the design of dissemination strategies. Social science & medicine (1982). 1999; 48:633–46. [PubMed: 10080364]
26. Carney TJWMMAMJJHDA. Organizational Factors Influencing the Use of Clinical Decision Support for Improving Cancer Screening Within Community Health Centers. International Journal of Healthcare Information Systems and Informatics. 2014; 9:1–29.
27. Carney TJMGPJJMAMWMWBHDA. Using computational modeling to assess the impact of clinical decision support on cancer screening improvement strategies within the community health centers. YJBIN Journal of Biomedical Informatics. 2014
28. Carley, KM., Reminga, J., Storrick, J., De Reno, M. ORA User's Guide 2009. Ft. Belvoir: Defense Technical Information Center; 2009. Carnegie-Mellon Univ Pittsburgh Pa Inst Of Software Research I.
29. Bruque S, Moyano J, Eisenberg J. Individual Adaptation to IT-Induced Change: The Role of Social Networks. Journal of Management Information Systems. 2008; 25:177–206.
30. Sargent RG. Validation and Verification of Simulation Models. WINTER SIMULATION CONFERENCE. 2004; 1:17–28.
31. Carley, KM. Validating Computational Models: Department of Social and Decision Sciences. Carnegie Mellon University; 1996.
32. Wolcott, HF. Transforming qualitative data : description, analysis, and interpretation. Thousand Oaks, Calif: Sage Publications; 1994.
33. Chen P-S, Gonyea R, Sarraf S, et al. Analyzing and interpreting NSSE data. New Directions for Institutional Research. 2009; 2009:35–54.

34. Merrill J, Bakken S, Rockoff M, Gebbie K, Carley KM. Description of a method to support public health information management: Organizational network analysis. Journal of biomedical informatics. 2007; 40:422. [PubMed: 17098480]

35. Wasserman, S., Faust, K. Social network analysis : methods and applications. Cambridge; New York: Cambridge University Press; 1994.

36. Axtell R, Axelrod R, Epstein JM, Cohen MD. Aligning simulation models: A case study and results. Comput Math Organiz Theor Computational and Mathematical Organization Theory. 1996; 1:123–41.

37. Thomas SD, Hathaway DK, Arheart KL. Face validity. Western journal of nursing research. 1992; 14:109–12. [PubMed: 1546473]

38. Schreiber C, Carley K. Going Beyond the Data: Empirical Validation Leading to Grounded Theory. Computational & Mathematical Organization Theory. 2004; 10:155–64.

39. Nemeth C, Connor M, Klock P, Cook R. Discovering Healthcare Cognition: The Use of Cognitive Artifacts to Reveal Cognitive Work. Organization Studies. 2006; 27:1011–35.

40. Niland JC, Rouse L, Stahl DC. An informatics blueprint for healthcare quality information systems. Journal of the American Medical Informatics Association : JAMIA. 2006:13.

41. Salas E, Rosen M, DiazGranados D. Expertise-Based Intuition and Decision Making in Organizations. Journal of Management. 2010; 36:941–73.

42. Anderson JA, Willson P. Knowledge management: organizing nursing care knowledge. Critical care nursing quarterly. 2009:32.

43. Sittig DF, Wright A, Simonaitis L, et al. The state of the art in clinical knowledge management: an inventory of tools and techniques. International journal of medical informatics. 2010; 79:44–57. [PubMed: 19828364]

13. Appendices

Appendix I: Facility-Level Cancer-Screening Performance Agents, Tasks, and Knowledge Elements and Their Assumptions

Patient-Care Agent Classification

Agent Categories:	Task Knowledge Impacting Performance is Informed by:	Knowledge Absorption (Homophily Knowledge) is Informed by:	Rationale and/or Assumptions:
Firm View – Patient Care	• Clinic Processes • Work Importance of Cancer-Screening Tasks • CDS and IS Practices • Delivery System Design for Cancer Screening (e.g. Role Responsibility, Overlap, and Clinical Champions) • Information-Dissemination Strategies	• Supportive Senior Leadership Environment • Supportive Local (Functional) Leadership Environment • Team Characteristics	• Human Agent (assumed 60% of firm staff) • % is arbitrary and not meant to represent any single firm within the sample • Patient-Care Agents are active in their ability to interact with other agents in the network • Leadership and Team interactions are viewed as opportunities for firm mission, goals, objectives, culture, and performance to be distributed

Agent Categories:	Task Knowledge Impacting Performance is Informed by:	Knowledge Absorption (Homophily Knowledge) is Informed by:	Rationale and/or Assumptions:
	• Provider IT-Performance Expectancy • Electronic-Information Retrieval and Availability		throughout the firm

Firm View – Administrative-Care Agent Classification

Agent Categories:	Task Knowledge Impacting Performance is Informed by:	Knowledge Absorption (Homophily Knowledge) is Informed by:	Rationale and/or Assumptions:
Firm View – Administrative Care	• Cancer-Screening Rate-Reporting Behavior (Provider Level) • Cancer-Screening Rate-Reporting Behavior (Facility Level) • Payer Mix (Insurance Type) – Uninsured Population – Medicare Population – Medicaid Population – Commercial Insurance Population • Financial Readiness (Cash Reserves) • Organizational Structure and Size • Information-Dissemination Strategies • Patient Demographics – Patient Age – Patient Language	• Supportive Senior Leadership Environment • Supportive Local (Functional) Leadership Environment • Team Characteristics	• Human Agent (assumed 40% of firm staff) • % is arbitrary and not meant to represent any single firm within the sample • Administrative Agents are active in their ability to interact with other agents in the network • Leadership and Team interactions are viewed as opportunities for firm mission, goals, objectives, culture, and performance to be distributed throughout the firm

IT-Systems Agent Classification – Clinical-Decision Support for Cancer Screening

Agent Categories:	Task Knowledge Impacting Performance is Informed by:	Knowledge Absorption (Homophily Knowledge) is Informed by:	Rationale and/or Assumptions:
IT Systems	• Cancer-Screening Rate-Reporting Behavior (Provider level) • Cancer-Screening Rate-Reporting Behavior (Facility Level) • Clinic Processes • Work Importance of Cancer-Screening Tests (CSTs) • Delivery-System Design for Cancer Screening (e.g. Role Responsibility, Overlap, and Clinical Champions)	• CDS Practices (IT Systems have all of this information)	• Non-human Agent • Specifically referencing IT in support of Cancer Screening • Assumes tie between cancer-screening performance and demand for IT-Systems Support IT Systems Activity is informed by: • Provider IT Performance Expectancy • Electronic Information Retrieval & Availability The % of this task knowledge that they have is based on: • EHR Functions and Capabilities

Outside-Collaborators Agent Classification

Agent Categories:	Task Knowledge Impacting Performance is Informed by:	Knowledge Absorption (Homophily Knowledge) is Informed by:	Rationale and/or Assumptions:
Outside Collaborators	• Cancer-Screening Rate-Reporting Behavior (Provider Level) • Cancer-Screening Rate-Reporting Behavior (Facility Level) • Clinic Processes	• No Explicit Homophily Knowledge sought for expertise (within the simulation)	• Assumes one-way communication of industry best practices to the firm • Scores represent the level of access and pace of infusion of this outside expertise Outside Collaborator Activity is informed by: • External Factors (e.g. Pressure, Support, Connectedness,

Agent Categories:	Task Knowledge Impacting Performance is Informed by:	Knowledge Absorption (Homophily Knowledge) is Informed by:	Rationale and/or Assumptions:
	• Work Importance of CSTs • Delivery System Design for Cancer Screening (e.g. Role Responsibility, Overlap, and Clinical Champions)		and Collaborative Agreements) • Environmental Assessment of Cancer-Screening Activities • Medical Specialist Availability

Cancer-Screening Test (CST) Agent Classification

Agent Categories:	Task Knowledge Impacting Performance is Informed by:	Knowledge Absorption (Homophily Knowledge) is Informed by:	Rationale and/or Assumptions:
Cancer-Screening Test	• Clinic Processes • Delivery System Design for Cancer Screening • CDS Practices • Information-Dissemination Strategies	• Work Importance of CSTs • Cancer-Screening Rate-Reporting Behavior Provider Level • Cancer-Screening Rate- Reporting Behavior Facility Level • Patient Demographics – Patient Age – Patient Language	• Non-human Agent • Agent is active all the time • Agent can be interacted with only by Patient-Care Agents • Agent cannot initiate interaction

Appendix II: Construct-TM Glossary

Summary Measure	Construct™ Coded Variable
X1 = HRSA Collaborative Experience	CollaborativeExp
X2 = Facility Age1–Number of Years receiving BPHC* funding	DateOpened
X3 = Facility Age2–Number of Years in any HRSA Collaborative	YrsHRSAFunded
X4 = Clinic Processes	ClinProcesses
X5 = Information Dissemination Strategies	InfoDissemination
X6 = Electronic Information Retrieval & Availability	ElecRetrieval
X7 = Electronic Health Record (EHR) Functions Capabilities	EHRFunctions
X8 = Work Importance of Cancer Screening Tasks	Screening_Task_Imp

Summary Measure	Construct™ Coded Variable
X9 = Cancer Screening Rate Reporting Behavior (Facility Level)	FacilityScreeningBehavior
X10 = Quality Improvement Strategies	QIStrategy
X11 = External Pressure, Support, Connectedness, and Collaborative Agreements	ExtAgreements
X12 = Delivery System Design for Cancer Screening (e.g., Role Responsibility, Overlap, and Clinical Champions)	SystemDesign
X13 = Supportive Senior Leadership Environment	SrLeadership
X14 = Supportive Local (Functional) Leadership Environment	LocalLeadership
X15 = Team Characteristics	Team
X16 = Medical Specialist Availability	MedSpec
X17 = Organizational Structure & Size	OrgSize
X18 = Financial Readiness1–Total Budget	Budget_Size
X19 = Financial Readiness2–Ratio of Revenues to Expenses	CashResearves
X20 = Payer Mix1–% Uninsured	UninsuredPop
X21 = Payer Mix2a–% Medicare	MedicarePop
X22 = Payer Mix2b–% Medicaid	MedicaidPop
X23 = Payer Mix2c–% Commercial Insurance	CommercialPop
X24 = Payer Mix2d–% Self Pay	SelfPayPop
X25 = Patient Demographics (Language)	PatientLanguage
X26 = Patient Demographics (Occupation Migrant Worker)	MigrantPop
X27 = Patient Demographics (Living Homeless)	HomelessPop
X28 = Patient Demographics (Age)	PatientAge
X29 = Environmental Assessment of Cancer Screening and Follow-up Activity via Provider Performance Feedback'	EnvAssessment
X30 = Cancer Screening Rate Reporting Behavior (Provider Level)'	ProviderScreeningBehavior
X31 = Provider IT Performance Expectancy	IT_Beliefs
Y1 = CDS Capacity for Measuring Cancer Screening (CDS1)	IT_Capacity
Y2 = Use of CDS Provider Prompts at Point-of-Care (CDS2)	Prompts
Y3 = Computerized Patient Reminders (CDS3)	Reminders
Y4 = Electronically Generated Correspondence with Results to Patients (CDS4)	PatientResults
YCDS = CDS Practices Score	CDS_score
YCSI = Cancer Screening Improvement Scores	Screening_rate

[*] Bureau of Primary Health Care

Appendix III: Task Definitions Used in Study

Descriptive	Index	Descriptive	Index
Budget	0	Medical Specialization	14
Cash Reserves	1	Medicare	15
CDS Score	2	Patient Age	16
Clinic Processes	3	Patient Language	17

Descriptive	Index	Descriptive	Index
Commercial Insurance	4	Patient Results	18
Electronic Retrieval	5	Provider Prompts at point-of-care	19
Environmental Assessment	6	Provider Screening Behavior	20
Facility Screening Behavior	7	Clinical Reminders	21
Info Dissemination	8	Screening Task	22
Insurance Type	9	Senior Leadership	23
IT Beliefs	10	System Design	24
IT Capacity	11	Team	25
Local Leadership	12	Uninsured	26
Medicaid	13		

Highlights

- We measure community health center cancer screening performance using dynamic network analysis and computational modeling
- Single point-in-time surveys on process improvement may lack information to project future success
- High performance organizations display differing network characteristics than lower performers
- Network structure reveals factors that shape diffusion and sustainability of best practices

Figure 1a

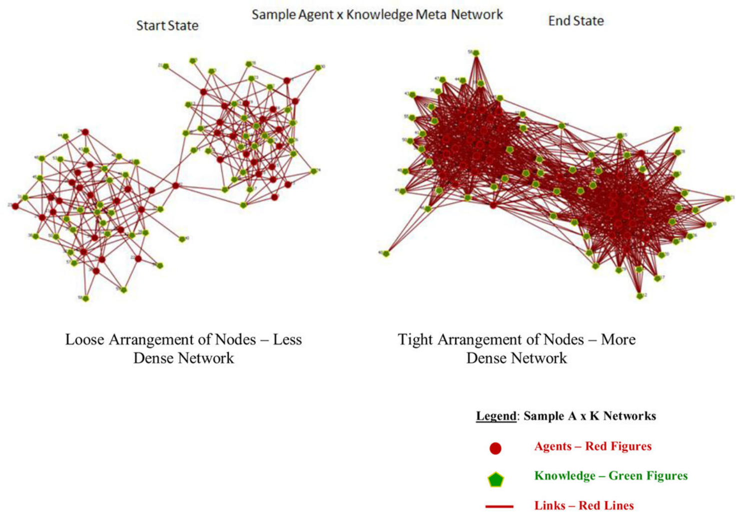
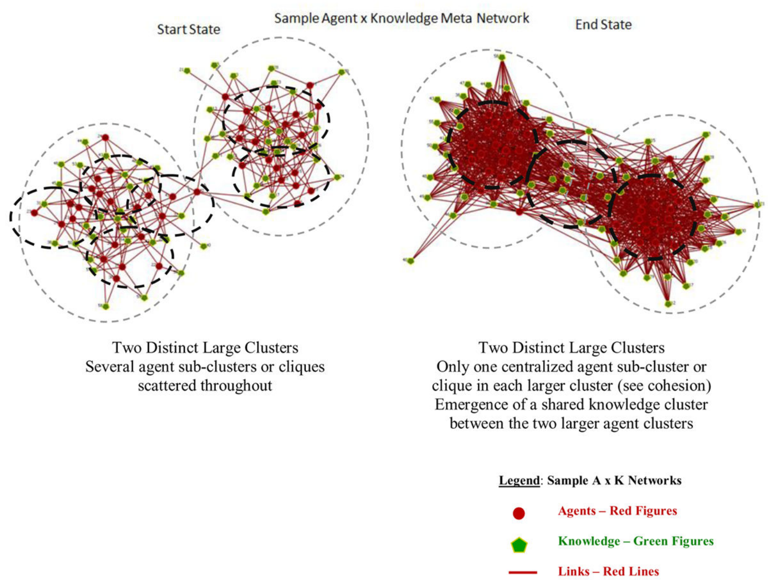


Figure 1b



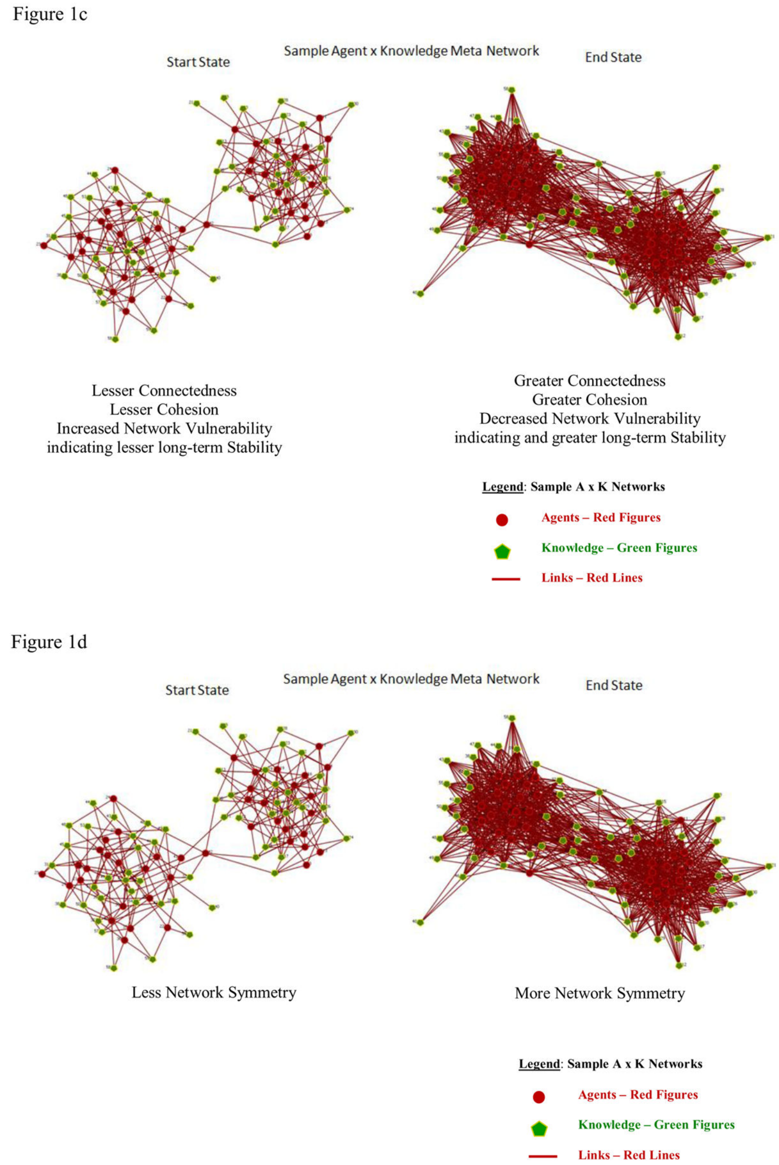


Figure 1.

Figure 1a: Sample A × K Network Examining Network Density

Figure 1b: Sample A × K Network Examining Network Clusters and Cliques

Figure 1c: Sample A × K Network Examining Network Connectedness and Cohesion

Figure 1d: Sample A × K Network Examining Network Symmetry

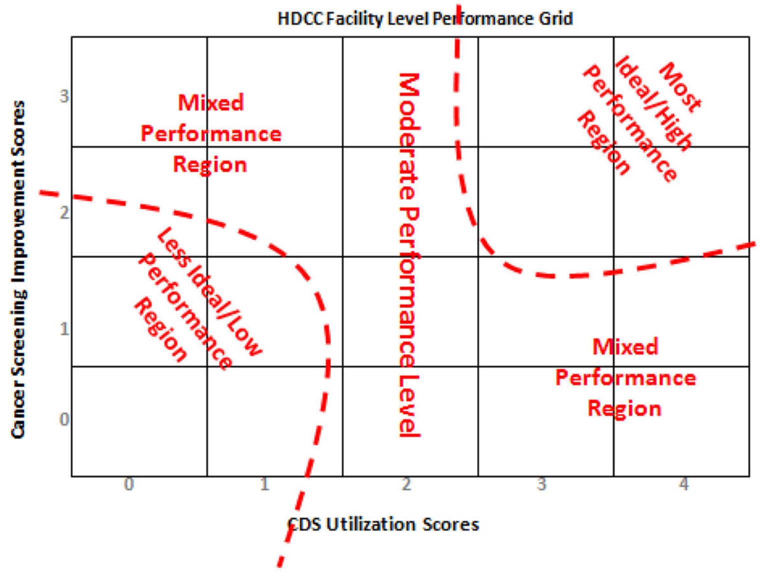


Figure 2.
CHCs Performance-level Matrix

Composite Performance Level Measure*	Position	Description	Frequency
High/High (HH)	Top Right	Ideal	24
Medium/High (MH)	Top Center	Moderate	5
Medium/Low (ML)**	Bottom Center	Moderate	1
High/Low (HL)	Top Left	Mixed	3
Low/High (LH)	Bottom Right	Mixed	6
Low/Low (LL)	Bottom Left	Less than Ideal	3

*No health centers were assigned to permutations not shown (e.g. HM, MM).
**Since only one health center occupied this level and had several missing data elements, ML was not included in the final analysis.

Figure 3.

Basis of Performance-level Assignment to Community Health Centers

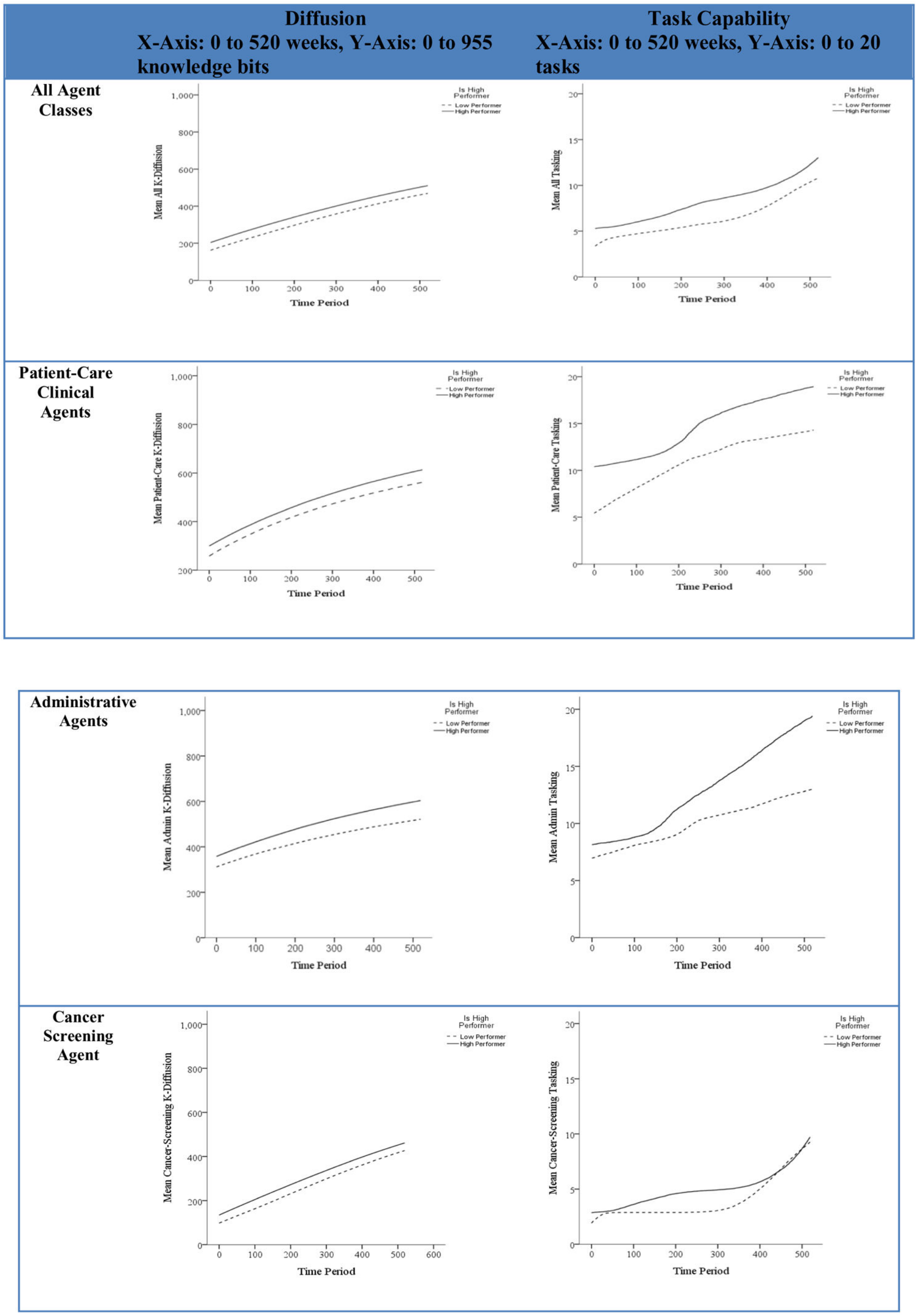


Figure 4.

Diffusion and Task Capability over time.

HH firms are solid, while LL firms are dashed.

Figure 5a

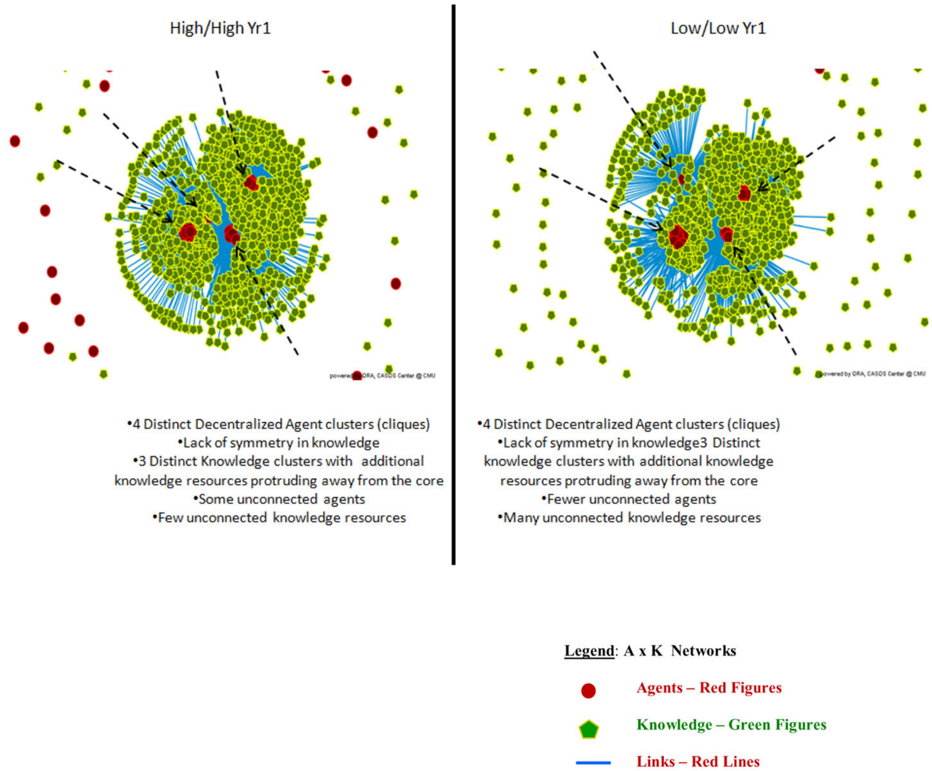

High/High Yr1
Low/Low Yr1
•4 Distinct Decentralized Agent clusters (cliques)
•Lack of symmetry in knowledge
•3 Distinct Knowledge clusters with additional knowledge resources protruding away from the core
•Some unconnected agents
•Few unconnected knowledge resources
•4 Distinct Decentralized Agent clusters (cliques)
•Lack of symmetry in knowledge3 Distinct knowledge clusters with additional knowledge resources protruding away from the core
•Fewer unconnected agents
•Many unconnected knowledge resources
Legend: A x K Networks
Agents – Red Figures
Knowledge – Green Figures
Links – Red Lines

Figure 5b

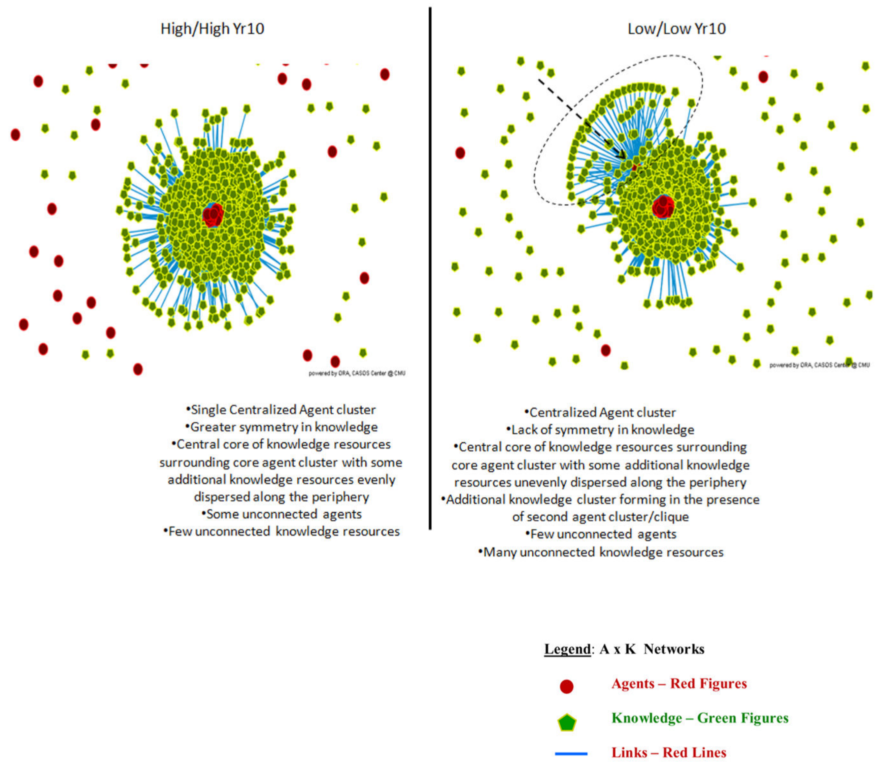


Figure 5.

Figure 5a: Agent-to-Knowledge (Knowledge Utilization) Network Configurations at Start of Simulation

Figure 5b: Agent-to-Knowledge (Knowledge Utilization) Network Configurations at the End of Simulation

Figure 6a

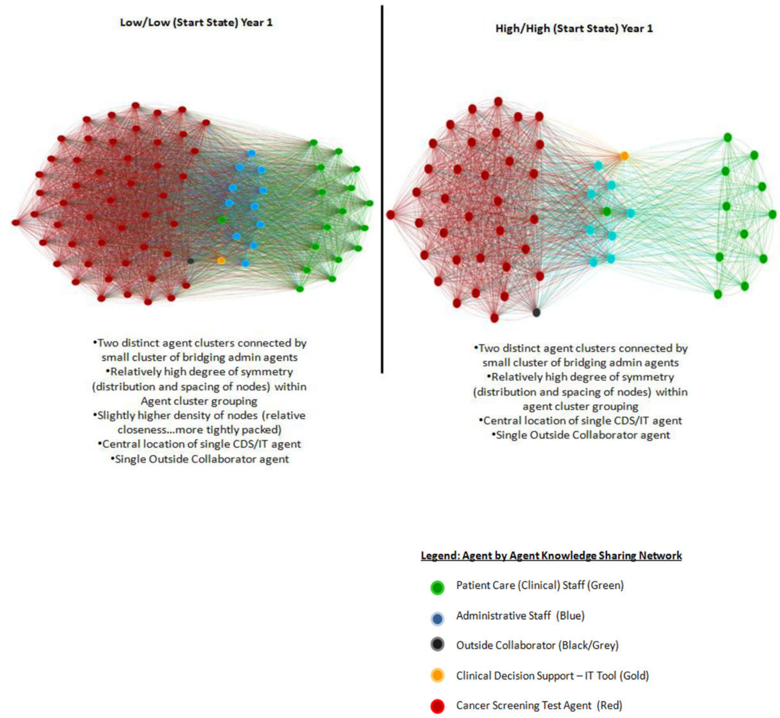


Figure 6b

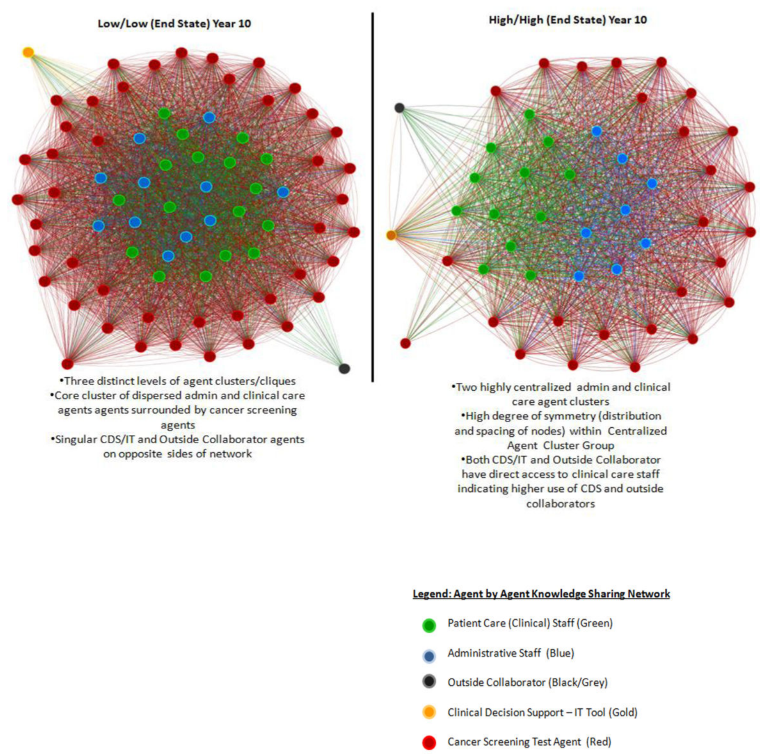


Figure 6.

Figure 6a: Agent-to-Agent (Knowledge-Sharing) Network Configurations at Start of Simulation.

Agents are colored by their membership groups.

Figure 6b: Agent-to-Agent (Knowledge-Sharing) Network Configurations at End of Simulation.

Agents are colored by their group memberships.

Table 1

Agent definitions are as follows:

Firm_Start	0
Firm_End	(100 × Financial Readiness_Budget) − 1
Patient Staff_Start	0
Patient Staff_End	0.6 × Firm_End
Adminstrative Staff_Start	Patient Staff_End + 1
Adminstrative Staff_End	Firm_End
IT System_Start	Firm_End + 1
IT System_End	IT System_Start
Outside Collaborators_Start	IT System_End + 1
Outside Collaborators_End	Outside Collaborators_Start
Cancer Screening Test_Start	Outside Collaborators_End + 1
Cancer Screening Test_End	Cancer Screening Test_Start + 3 * (Patient_Staff_End)

Table 2

System-wide Schema for All Knowledge

Start_index	End_Index	Construct-TM Name	Descriptive Variable Name
0	49	SrLeadership	Supportive Senior Leadership Environment
50	99	LocalLeadership	Supportive Local Leadership Environment
100	149	Team	Team Characteristics
150	199	ClinProcesses	Clinic Processes
200	249	Screening_Task_Imp	Work Importance of Cancer Screening Tasks
250	297	CDS_score *	CDS Practices (IT Capacity, Prompts, Reminders, Patient Results)
298	347	SystemDesign	Delivery System Design for Cancer Screening
348	397	ProviderScreeningBehavior	Cancer Screening Rate Reporting Behavior Provider Level
398	445	InsuranceType *	Uninsured, Public-Medicaid, Pubic-Medicare, Commercial, Self-pay
446	495	CashReserves	Financial Readiness_RevenueToExpense
496	545	Budget_Size	Combined Size and Budget
546	595	InfoDissemination	Information Dissemination Strategies
596	645	PatientAge	Patient Demographics (Age)
646	695	PatientLanguage	Patient Demographics (language)
696	745	ITBeliefs	Provider IT Performance Expectancy
746	795	ElectronicRetrieval	Electronic Information Retrieval & Availability
796	845	EnvironmentalAssessment	Environmental Assessment of Cancer Screening and Follow-up Activity via Provider Performance Feedback
846	895	MedicalSpecialization	Medical Specialist Availability
896	945	FacilityScreeningBehavior	Cancer Screening Rate Reporting Behavior Facility Level

* These knowledge buckets are made of four separate sub-indices, each with 48 bits rather than 50.

Table 3

Mean and SD by HDCC Performance Grouping

	Facility Categories (Based on Cancer Screening and Clinical Decision Support Scores)			
	Low/Low (LL)		High/High (HH)	
Summary Measures	**Mean**	**SD**	**Mean**	**SD**
HRSA Collaborative Experience	2.33	0.58	2.96	1.12
Facility Age (1): Number of Years Receiving BPHC* funding	13.50	2.12	22.86	10.07
Facility Age (2): Number of Years in any HRSA Collaborative	12.00	0.00	17.85	10.56
Clinic Processes	2.33	0.58	2.83	0.92
Information Dissemination Strategies	17.00	4.36	17.38	2.89
Electronic Information Retrieval & Availability	1.00	0.00	0.67	0.70
Electronic Health Record (EHR) Functions Capabilities	1.00	1.73	5.08	1.98
Work Importance of Cancer-Screening Tasks	23.33	0.58	24.67	0.76
Cancer-Screening Rate Reporting Behavior (Facility Level)	3.33	2.52	4.21	1.86
Quality-Improvement Strategies	34.00	2.00	31.92	8.67
External Pressure, Support, Connectedness, and Collaborative Agreements	34.00	2.00	31.92	8.67
Delivery-System Design for Cancer Screening (e.g. Role Responsibility, Overlap, and Clinical Champions)	54.33	31.66	68.92	9.99
Supportive Senior Leadership Environment	25.67	3.21	25.83	3.75
Supportive Local (Functional) Leadership Environment	13.33	1.53	13.71	1.88
Team Characteristics	38.33	5.13	36.63	3.55
Medical Specialist Availability	6.67	5.77	7.46	4.10
Organizational Structure & Size	25.67	26.01	47.71	47.66
Financial Readiness (1)	$14,486,121	$7,797,802	$11,115,815	$10,834,573
Financial Readiness (2)	5.00	0.00	4.41	1.18
% Uninsured	12.50	10.61	36.41	17.20
% Medicare	5.50	4.95	16.59	18.22
% Medicaid	74.50	14.85	44.77	22.56
% Commercial Insurance	4.50	3.54	10.41	8.26
% Self-Pay	15.50	6.36	28.29	21.43

	Facility Categories (Based on Cancer Screening and Clinical Decision Support Scores)			
	Low/Low (LL)		High/High (HH)	
Summary Measures	**Mean**	**SD**	**Mean**	**SD**
Patient Demographics (Language)	6.00	5.66	19.14	22.22
Patient Demographics (Occupation: Migrant Worker)	1.00	1.41	1.74	4.82
Patient Demographics (Living: Homeless)	0.00	0.00	2.47	3.22
Patient Demographics (Age)	1.00	0.00	1.62	0.80
Environmental Assessment of Cancer Screening and Follow-up Activity via Provider Performance Feedback	52.67	9.87	57.33	7.18
Cancer-Screening Rate Reporting Behavior (Provider Level)	5.67	0.58	5.54	1.06
Provider IT-Performance Expectancy	27.00	4.00	25.13	3.94
CDS Practice Scores	0.67	0.58	3.42	0.50
Cancer-Screening Improvement Scores	0.33	0.58	2.71	0.46

Indicate areas where LL CHCs score highest

Indicate areas where HH CHCs scored highest

Indicate areas where LL and HH CHCs relatively even

Note: Medium/Low (ML) has been omitted for consideration within this simulation due to an N of 1 and several missing data elements.

Table 4

Descriptive Statistics for Displaying Numerical Characteristics by Performance Level

	LL			HH			
Network-Level Measure	**Mean**	**Std. Dev**	**Median**	**Mean**	**Std. Dev**	**Median**	**Sig.**
Change in Density*	1.75	0.016	1.75	1.21	0.014	1.21	< 0.001
Change in In-Degree*	−0.13	0.016	−0.13	−0.09	0.017	−0.09	< 0.001
Change in Row Breadth**	0.01	0.002	0.01	0.02	0.004	0.02	< 0.001
Change in Column Breadth^	0.08	0.011	0.09	0.07	0.007	0.07	< 0.001
Column Redundancy^	2.21	0.08	2.19	1.46	0.075	1.46	< 0.001
Change in Link Count^^	1.93	0.018	1.93	1.58	0.023	1.58	< 0.001
Change in Tasking^^	2.21	0.08	2.19	1.46	0.08	1.46	< 0.001

* (Link Count, Density, Average in Degree): Low performers actually have many connections, and interaction is not stifled

** (Row Breadth): High performers are more specialized, and only become more so

^ (Column Breadth and Column Redundancy): High performers are connected more strongly to their connections and have fewer redundant assignments of knowledge.

^^ (Measures of the performance of the organization): Low-performing organizations improve rapidly, but are not making up lost ground due to relative starting points. Figure 4 illustrates in more detail.