DEVELOPMENT OF DASHBOARDS TO PROVIDE FEEDBACK TO HOME CARE NURSES

Principal Investigator: Dawn Dowding,[1,2] PhD, RN, FAAN

Team Members:
David Russell,[1] PhD
Jacqueline Merrill,[2] PhD, MPH, RN, FAAN, FACMI
Yolanda Barrón,[1] MS
Nicole Onorato,[1] BS
Robert Rosati,[3] PhD
Karyn Jonas,[1] BS, RN

Organization:
1. Center for Home Care Policy and Research, Visiting Nurse Service of New York, New York, NY 10001
2. Columbia University School of Nursing, New York, NY 10032
3. VNA Health Group, Red Bank, NJ 07701

Dates: 09/30/2015 – 03/31/2017

Federal Project Officer: Tiffani J. Bright, PhD

This project was supported by grant number R21HS023855 from the Agency for Healthcare Research and Quality. The content is solely the responsibility of the authors and does not necessarily represent the official views of the Agency for Healthcare Research and Quality.

Grant Award Number: 1R21HS023855

ABSTRACT

Purpose: To design dashboards that could provide feedback on metrics related to the care of patients with congestive heart failure (CHF) to nurses working in a home care setting and to evaluate their usability.

Scope: The study was conducted primarily in one home healthcare agency in the north east of the US, with additional survey data collected from a second agency in the same region. The study explored information that nurses would like to see provided in visualized format and individual differences in comprehension of data presented in different visualized formats. Finally a prototype dashboard, designed for use by homecare nurses at the point of care, was evaluated for usability.

Methods: Focus groups with nurses and other home care professionals (N=61), a multi-factorial experimental design using an online survey (n=195) and usability evaluation (n=32) using rapid feedback, heuristic evaluation with experts (n=3) and formal usability testing.

Results: Nurses prioritized information about a patient's signs and symptoms (specifically increase in weight) as being an area of practice where they would like to receive feedback. This should be in 'real time' and provide data on trends over time. Individual differences in the ability to understand information presented in different graphical formats were identified; information presented in the format of a bar graph was associated with the highest comprehension levels, particularly for individuals with low numeracy and low graph literacy. The dashboard prototype had high usability (mean SUS = 73.2, SD 18.8) and was evaluated positively be nurse users.

Key Words: Clinical dashboard, graph literacy, numeracy, nursing informatics, home health

1. PURPOSE

The goal of this pilot and feasibility study was to design dashboards that could provide feedback on metrics related to the care of patients with congestive heart failure (CHF) to nurses working in a home care setting and evaluate their usability. The study addressed the following aims:

Aim 1: Operationalize quality performance metrics related to the management of patients with CHF that can be used by home care nurses to improve their practice.
Aim 2: Provide feedback to home care nurses which maximizes their comprehension of the data.
Aim 3: Develop a set of flexible, dynamic quality dashboards individualized to nurse user preferences.

2. SCOPE (Background, Context, Settings, Participants, Incidence, Prevalence)

2.1. Background

The quality and safety of health care has been the focus of a number of quality improvement strategies in health care organizations.[1] In 2014 it is estimated that approximately 4.9 million individuals in the US were cared for in a home health setting.[2] Home healthcare agencies (HHAs) in the US typically provide care to patients post discharge from a hospital, where they may have received skilled nursing care, rehabilitation therapies (e.g. physical and occupational therapy) and social work support. Hospital readmission of home care patients represents an indicator of poor care quality, and Medicare beneficiaries who have a diagnosis of CHF have the highest overall re-hospitalization rates across all health care settings.[3] A recent study focusing on national data for HHAs identified an increased risk of hospital readmission for patients with CHF as the primary diagnosis (adjusted odds ratio =1.42, 95% CI 1.37-1.48).[4] Although many risk factors that have been identified as predictors of a patient's likelihood of re-hospitalization are not modifiable by clinicians (e.g., race/ethnicity, gender), there are others which are under their direct control. For example, for CHF patients it is recommended that monitoring should include medication reviews and regular assessments of fluid status (including weight gain) nutritional status and functional capacity.[5 6]

A systematic review of quality improvement strategies suggests that the provision of feedback at an organizational level does not have any discernable effect on the behavior or performance of individual clinicians or patient outcomes.[7] In contrast, providing feedback to clinicians on individual performance has been shown to lead to potentially important improvements in practice, although it is unclear precisely what methods of feedback lead to improvements.[1] One method of providing feedback to individual clinicians are dashboards; a form of health information technology (HIT) that use data visualization techniques to provide feedback to individuals on their performance compared to quality metrics.[8] In comparison to other forms of feedback, dashboards have the potential to provide feedback in 'real time' when clinicians are engaged in care activities, rather than providing a retrospective summary of performance.[9] The use of data visualization (using graphical displays to summarize data) is thought to improve comprehension[10 11] and reduce cognitive load,[12] leading to

more effective decision making. However, clinicians' ability to comprehend numerical and graphical data tends to vary by their level of experience and expertise.[13 14] There is also evidence to suggest that cognitive factors such as spatial ability affect how individuals perceive and use data presented in visualizations.[15] Technology systems are more likely to be used effectively when they are designed with user cognition in mind and enable clinicians to make modifications that fit their existing work practices.[16-18]

In this study we proposed to build upon existing research evidence to design dashboards to provide feedback to nurses at the point of care, on key quality metrics related to the care of patients with CHF. This included using existing guidelines to identify priority areas for developing feedback information (Aim 1) and exploring the impact of individual nurses' levels of numeracy (the ability to understand numerical information) and graph literacy (the ability to understand information presented graphically)[14 19-22] on their ability to understand information presented in a variety of visualized formats (Aim 2). The results of these two elements of the study were then used to develop dashboards that were designed to be individualized to user characteristics and preferences, and integrated into existing technology systems. The prototype dashboard was then evaluated for usability (Aim 3).

2.2. Setting

The study was conducted predominantly at one HHA; the VNSNY which operates a large not-for-profit home health care agency serving a diverse population of individuals in the New York metro area. In addition the VNA Health Group in New Jersey, the largest not-for-profit provider of home health and community-based care in New Jersey, participated in the survey element of the study (Aim 2).

2.3. Participants

Aim 1 (focus groups): 61 participants from 1 HHA in 6 focus groups (53 nurses, 6 physiotherapists, 1 occupational therapist and 1 speech and language therapist). We identified potential home care teams to participate in the focus groups at random and then approached the manager of each team to ask permission to organize a focus group either pre or post one of their team meetings. The selected team members were then invited to attend a focus group at a specified time. It was made clear that participation was voluntary.

Aim 2 (Online survey): 195 participants (all nurses) from the 2 HHAs (129 from Agency 1 and 66 from Agency 2). All potential participants (i.e. all RNs identified by the two agency Human Resources departments as working in the agency) were sent an email describing the purpose of the study and containing a live link to the survey website. Eligibility criteria included: 1) Registered Nurses (RN) who were currently employed at a Certified Home Health Agency (CHHA), and 2) visiting patients once or more per week.

Aim 3 (Usability Evaluation): 32 home care nurses working in 1 HHA (2 nurses as co-designers, 10 provided rapid feedback, 20 participated in the usability evaluation). Nurses were volunteers recruited via email (co-designers and usability evaluation) or after team meetings (rapid feedback). In addition, 3 experts were recruited to conduct a

heuristic evaluation of the dashboard prototype. Criteria for selection as an expert were that they were a nurse, had published in the field of informatics and had some expertise in the field of data visualization.

3. METHODS (study design, data sources/collection, interventions, measures, limitations)

3.1. Study Design

Aim 1: Focus groups were used to identify existing quality performance measures related to the care of patients with CHF that are *relevant* to home care nurses, and that are *actionable* (i.e. under the control of the home care nurse), from the nurses' viewpoint.

Aim 2: A multi-factorial experimental design, with data collected via an on-line survey. The survey was used to explore the impact of numeracy and graph literacy of home care nurses on their comprehension of visualized data.

Aim 3: The results of Aims 1 & 2 were used to develop the prototype dashboards using a design science research approach consisting of 3 iterative processes:

i) Contextual inquiry was used to observe potential users of the system and interviews to explore how they do their work;
ii) Prototype systems were built and evaluated in an iterative cycle using rapid feedback from a sample of nurses;
iii) Usability evaluation of the final dashboard design. The dashboard was evaluated for function (i.e. how useful is it?), users (i.e. how satisfying do the users find it?), representations and task (i.e. how usable is it?).

3.2. Data Sources/Collection

3.2.1. Aim 1 (Focus Groups)

Focus groups took place between November 2015 and February 2016. Prior to the focus groups, the research team extracted statements from the evidence-based guidelines used to develop agency care protocols.[23-27] In total this exercise identified 28 statements used in the focus groups. Each statement was placed on a poster; at the beginning of each focus group the posters were attached to the interview room wall and participants were asked to identify the top five evidence based care statements that would be a priority for them to manage during a home care visit for a patient with heart failure. They were provided with color-coded stickers that also had a number (1= top priority through to 5 = least in the top 5) and attached them to the relevant posters. The subsequent focus group discussion was structured around the statements most important to participants (the highest priority ranking). The focus group guide consisted of general questions to explore the participants' experiences providing the care related to each statement, and technology that might be developed to help support their care. All interviews were audio-taped and professionally transcribed.

3.2.2: Aim 2 (on-line survey)

All potential participants were sent an email describing the purpose of the study and containing a live link to the survey website. Each link was unique to the user, allowing

them to save and return to complete the survey at a later date. Reminders were sent to all non-responders two, three and four weeks after the initial email. On completion nurses could provide their contact details (separate from survey responses) to receive a $20 gift voucher in recognition of their time. Questions to capture demographic characteristics and questions to measure numeracy, graph literacy and comprehension (see section 3.4.) were combined into a single survey instrument (which could be accessed/completed on a mobile/tablet device as well as a personal computer) with a total of 58 items, and formatted using an online software program (www.surveygizmo.com). Participants firstly completed screening questions to ensure their eligibility to participate in the study and asked to complete a demographic questionnaire followed by both the graph literacy scale and numeracy scale. They were then randomized (using the logic programming from the survey software) into one of four experimental groups (Table 1). The study was designed so that each participant received data for each quality indicator in one of each different format of data presentation. This method ensured that any differences observed in comprehension across graph formats was related to differences in numeracy or graph literacy, and not due to the type of data displayed or to the characteristics of nurses.

Experimental Group	**Quality Indicator 1**	**Quality Indicator 2**	**Quality Indicator 3**	**Quality Indicator 4**
1	Table	Line Graph	Bar Chart	Spider Graph
2	Line Graph	Bar Chart	Spider Graph	Table
3	Bar Chart	Spider Graph	Table	Line Graph
4	Spider Graph	Table	Line Graph	Bar Chart

Table 1: Experimental Group Conditions

3.2.3. Aim 3 (Usability Evaluation)

Results from the focus group study and experimental study identified what areas of practice associated with the care of CHF patients' home care nurses wish to receive as feedback. In order to ensure the dashboards were integrated into the nurse workflow, we used a data gathering approach known as contextual inquiry,[28 29] which combines observations of individuals who are the potential users of a system (in this case home care nurses) in their work environment and interviews with those individuals to explore how they do that work. A researcher accompanied the nurse to at least one patient visit where they observed them as they carried out care, and then asked questions about their observations afterwards. The researcher took notes to inform their interviews. We used a defined process which included verifying with the nurses that the researcher's understanding of the work they were observing was correct and asking for nurses' views and comments.

Based on the observations and interview feedback a set of prototype dashboards for the monitoring of patient vital signs (weight, blood pressure, temperature, pulse, respirations and pain score) were developed. The dashboards were firstly developed as paper based prototypes and then in a visual design (using an interactive web based interface). The paper prototypes were formatted into posters and placed on poster boards. Nurses were asked to provide written consent to participate in the study, and

then asked to write down their thoughts/feedback on the prototype design (including layout, colors, information presented in the dashboard) and suggestions for improvements. In general, the feedback sessions lasted between 5-10 minutes. Following this feedback an interactive web-based prototype was developed.

The interactive web-based prototype dashboard was evaluated for function (how useful is it?), users (how satisfying do the users find it?), representations and task (how usable is it?). We arranged to meet all the nurses who participated in the usability evaluation at one of the VNSNY regional offices at a time that was convenient for them (in liaison with the VNSNY regional management team). Nurses were asked to provide written informed consent to participate in the study. They were then given a written series of tasks to complete using the web based dashboard prototype. The tasks required them to extract data from the dashboard as well as explore its functionality (such as switching between different graph types, selecting different data to display and moving between the front patient case screen and the dashboard located in the patient's notes for vital signs recording). The nurses' interactions with the dashboard were recorded using Morae® software (which included recording the time they started and finished the tasks). Once they had completed the tasks, they were asked to complete the SUS and QUIS scales and provided verbal feedback on the dashboard and their experience of using it.

Experts for the heuristic evaluation were recruited via email, sent out to experts in data visualization and nursing informatics at Columbia University Medical Center. They were provided with an extended task list designed to enable them to explore the full functionality of the dashboard and the heuristic evaluation checklist. At a mutually agreed date and time, they used the task list to examine the dashboard, completed the evaluation checklist and provided written feedback on the dashboard design.

3.2. Measures

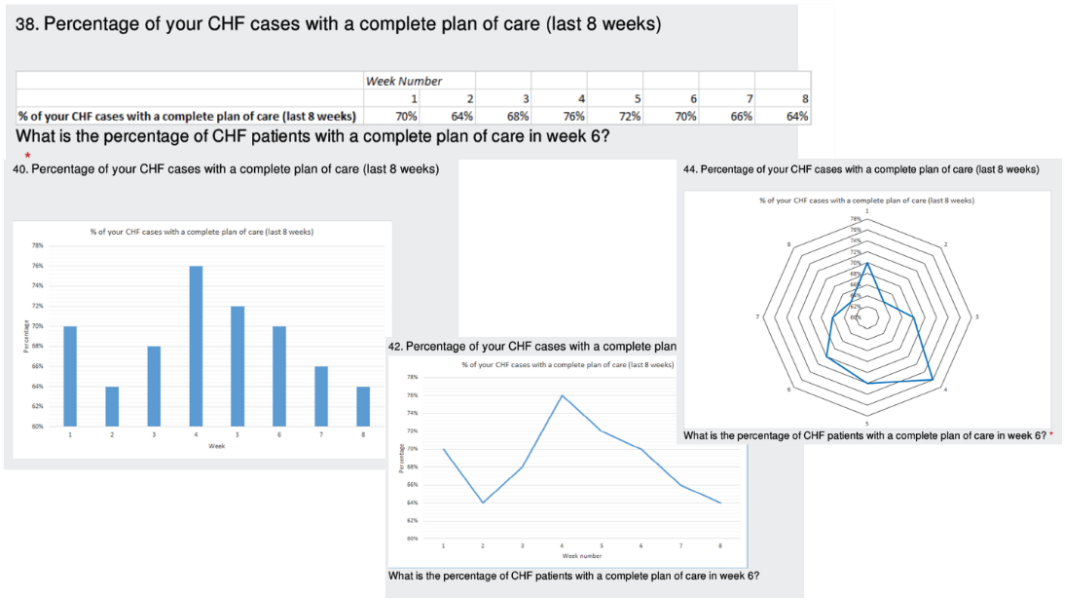
3.2.1. Aim 2 (Online Survey)

Numeracy: Numeracy was measured using the expanded numeracy scale.[30] The scale has 3 general numeracy items and 8 expanded items that frame questions in the context of health risks. Higher scores on the scale are associated with higher numeracy. The scale has been widely used as a method for assessing objective numerical competence, with high reliability (Cronbach α ranging from 0.7 – 0.75). Using a median cut-off individuals are categorized as having low numeracy if their score is less than 8, and high numeracy if their score is ≥ 8.[30]

Graph Literacy: Graph literacy was measured using a scale developed by Galesic and Garcia-Retamero[31] specifically for the health domain, to measure both basic and advanced graph reading skills and comprehension across different types of graphs.[31] The scale was originally tested in two samples from Germany and the US, with high internal consistency (α = 0.74 and 0.79 respectively).[31] A lower score on the scale represents lower graph literacy; using a median cut-off of 9, individuals with a score lower than 9 are considered to have low graph literacy and those ≥9 high graph literacy.

Data and Graph Comprehension: Four indicators representing feedback about the quality of care provided to heart failure patients (percentage of your CHF patients with a complete plan of care (last 8 weeks); percentage of high risk CHF patients seen within 24 hours of admission to home care; percentage of CHF patients with high risk of hospitalization; patient daily weight record) were developed. The data for each quality indicator could be presented in 1 of 4 different presentation formats (bar graph, line graph, spider/radar graph or table/text) (Figure 1). Following presentation of each quality indicator participants were asked questions regarding their comprehension of the data presented in the graph/table. In addition all participants were asked to complete questions to assess their comprehension of quality targets using four questions that required them to interpret data presented in a table and graph relative to a quality target.

Figure 1: Presentation formats for quality indicator information



3.3.2: Aim 3 (Usability Evaluation)

Usefulness was assessed through observation of nurse users as they interacted with the dashboard, comparing how they used the system with the functionality built into the dashboard. Specially designed usability software (Morae®)[32] was used to facilitate the recording of nurse interactions with the dashboard interface including verbalizations and on-screen activity (including mouse and keyboard inputs).

Satisfying: This is a measure of how useful the users find the system, and was measured using two established questionnaires, the system usability scale (SUS)[32] and the Questionnaire for User Interaction Satisfaction (QUIS).[33 34] The SUS is a flexible questionnaire designed to assess any technology and is relatively quick and easy to complete. It consists of 10 statements that are scored on a 5-point scale of strength of agreement, with final scores ranging from 0-100. A higher score indicates better usability. As a general rule, a system that has a score above 70 has acceptable

usability; a lower score means that the system needs more scrutiny and continued improvement.

The QUIS has been developed by a group of researchers at the University of Maryland and is specifically designed to measure user satisfaction with various components of a technology system which includes both overall system satisfaction and specific interface factors such as screen design and terminology used in the system.[33] [34] The questionnaire can be configured so that it fits the needs for user interface analysis. It has extremely high reliability (Cronbach's alpha = 0.95)[34] and has been used successfully for user evaluations across a variety of different technology systems in healthcare.[35-38] The QUIS divides overall usability into subcomponents, with each subcomponent consisting of a set of questions that can be used individually or with other subcomponents. For this study we used the QUIS short form, with the subcomponents part 3 (overall user reactions), part 4 (screen), part 5 (terminology and dashboard information, questions 1 and 2), part 6 (Learning, questions 1, 3 and 4) and Part 7 (system capabilities).

Usability: This was evaluated using heuristic evaluation and task analysis.[39] Heuristic evaluation involves asking a group of evaluators to independently assess an application, comparing it to a set of recognized usability principles (heuristics). The evaluators assess whether or not the heuristic has been violated and gives it a severity rating with 1 (a cosmetic problem only) through to 4 (usability catastrophe- imperative to be fixed before a product can be released). Each evaluator was asked to use the system and complete a heuristic evaluation form (specifically designed for this study based on established usability heuristics).

Task analysis was used to compare user performance in terms of time on task. We estimated the performance level of experts using the dashboard, by calculating the total time it took them to complete a set of tasks. The actual time that users spent using the dashboard was compared to the expert time. The data used to conduct the task analysis was recorded automatically during the nurse's use of the dashboard in the usability evaluation session by Morae® software.

3.3: Analytic Approach

3.3.1: Aim 1 (focus groups): Focus group transcripts were entered into a qualitative analysis program (NVivo 11). Firstly, all researchers read the transcripts of 2 interviews, to identify an initial set of broad themes related to the data and study aims. Next, two researchers (DD and JM) used the thematic analysis to develop a detailed coding framework, which they then applied to the two transcripts independently. Following this process, they met and discussed their initial coding, and reached consensus on a revised set of codes. This set of codes was used to categorize data across the interview transcripts and inter-rater reliability was calculated (Kappa = 0.84). Finally, all researchers reviewed the coded data and reached consensus on the findings that emerged from the data.

3.3.2: Aim 2 (online survey): Descriptive statistics were used to examine the demographic characteristics of the sample population, and chi-square or t-test were used to compare these characteristics between the two agencies. Mixed linear models were used to explore the impact of different display types on data comprehension, and to separately estimate the interaction effect between both high-low numeracy or high-low graph literacy on data comprehension. In addition a mixed linear model was used to estimate the interaction effects between the 4 groups of numeracy and graph literacy and type of display on data comprehension. Analyses were conducted separately for each agency as well as for all participants. Results were adjusted for multiple comparisons. All data were analyzed using SAS (version 9.4)

3.3.3: Aim 3 (usability evaluation): Usefulness: For each task conducted as part of the usability evaluation, we identified whether or not the participants completed the task with ease, with difficulty (requiring help to complete the task) or failed to complete it. The audio recordings of nurse interactions with the dashboard and their feedback was analyzed using content analysis.
Satisfying: The SUS was scored by converting responses to a scale with values from 0 to 4 (where 4 is the most positive response). The converted responses were added together and then multiplied by 2.5 as per the scoring instructions, giving a range of possible values from 0 to 100. Descriptive statistics were used to summarize the SUS scores across all evaluators of the system. The mean and standard deviation for each item on the QUIS was calculated and then graphically displayed providing an overall profile to identify areas that participants identified as being particularly good or bad. Usability: The output from a heuristic evaluation is a summary list of usability problems identified by the group of evaluators. For each usability principle (n=10) there were a list of usability factors (n=49). If the factor was present then the evaluator gave a score of 1 (Yes) and if it was not present, a score of 0 (No). We calculated scores for each heuristic by calculating the total number of factors (points) awarded, divided by the total number available. The higher the score, the greater the usability of the system.

The tasks associated with operating the dashboard were identified from the task list. Using the task list created for the usability evaluation we recorded 3 expert users (separate from those individuals participating in the heuristic evaluation) interacting with the dashboard and calculated the total time to complete the task list. We calculated the mean time for an expert to complete the tasks which was then used as our expert model. For each participant in the usability evaluation we calculated their total time using the dashboard. Descriptive statistics were used to summarize this data, including the average, maximum and minimum times taken to complete the task list, and how they compared to the expert model. We also recorded if users could accurately extract information from the dashboard.

4. RESULTS (Principal Findings, Outcomes, Discussion, Conclusions, Significance, Implications)

4.1. Principal Findings

4.4.1. Aim 1 (focus groups): Nurses prioritized information about a patients signs and symptoms (specifically weight and vital signs) and educational goal achievements as being areas of practice where a) they would find it useful to receive feedback and b) where they felt they could have an impact on outcomes. Discussion in the focus groups highlighted that they wanted this feedback in 'real time' (i.e. available when they opened a patient's note in the electronic health record) and for it to show trends over time. Nurses believe that having an easily accessible display for this information in a visualized format (e.g. a line graph, identification of which goals have been achieved) would enable them to plan care more effectively during a home care visit.

4.4.2. Aim 2 (online survey): The mean numeracy score for the total sample was 8.4 (SD 2.0) and numeracy was 9.7 (SD 2.4). Overall, nurses comprehended information in bar graphs more easily than other formats; mean 88% correct responses compared to 81 % correct responses for a table, 77% correct responses for a line graph and 41% correct responses for a spider/radar graph. Nurses were categorized as having either high (N=147, 75%) or low (N=48, 25%) numeracy and high (N=146, 75%) or low graph literacy (N=49, 25%). Numeracy and graph literacy were also related to data comprehension; nurses with high numeracy had higher comprehension (75% correct responses compared to 65% for low numeracy; -10.0 (SE 3.6) p=0.006) and nurses with high graph literacy had higher comprehension (77% correct responses compared to 63%; -14.4 (SE 3.6), p<0.001). There were significant interactions between numeracy and graph literacy and data presentation type; individuals with low numeracy or graph literacy had lower comprehension of information in spider/radar graphs. In addition individuals with low graph literacy were also less likely to comprehend data in a table (Table 2).

Numeracy/Graph Literacy Profile	**Percent point difference (SE)**	**p-value**
Numeracy (Low vs High)		
Line Graph	-7.0 (6.0)	0.27
Spider/Radar Graph	-26.0 (6.0)	**<0.01****
Table	4.0 (7.0)	0.53
Bar Graph	-11.0 (6.0)	0.09
Graph Literacy (Low vs High)		
Line Graph	-11.0 (6.0)	0.1
Spider/Radar Graph	-22.0 (6.0)	**<0.01****
Table	-22.0 (7.0)	**<0.01****
Bar Graph	-2.0 (6.0)	0.78

*p < 0.05 ** p<0.01

Table 2: Interaction between numeracy, graph literacy and display type

Nurses were then categorized into 4 categories; i) low numeracy and low graph literacy (n=24; 12%), ii) high numeracy and low graph literacy (n=24; 12%), iii) low numeracy and high graph literacy (n=25; 13%) and iv) high numeracy and high graph literacy (n=122; 63%). Nurses who had high scores on both numeracy and literacy had higher

correct responses; 81% compared to 78% for low numeracy and high graph literacy, 74% for high numeracy and low graph literacy and 54% for low numeracy and low graph literacy. There were also significant interaction effects depending on an individual's graph literacy and numeracy (Figure 2). Overall nurses with both low numeracy and graph literacy were less able to comprehend information across all display types. High graph literacy was associated with greater comprehension of spider/radar graphs and tables.

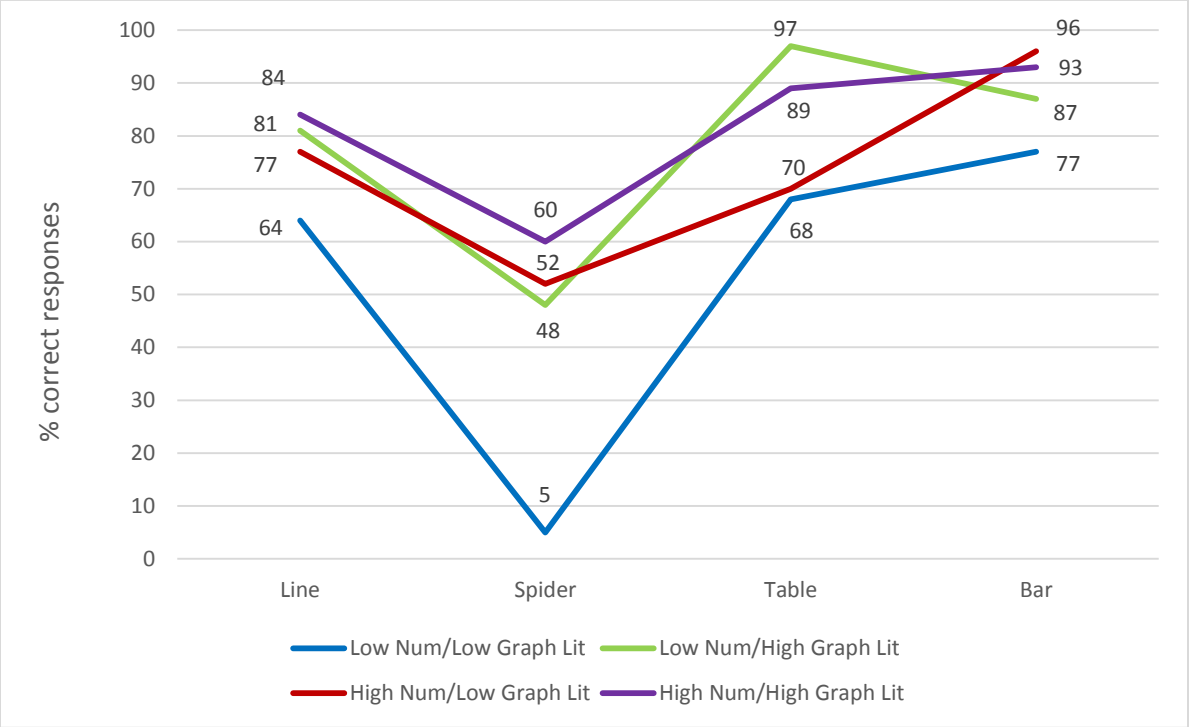


Figure 2: Interaction of graph literacy/numeracy and display type

4.4.3. Aim 3 (Usability Evaluation): Dashboard design: The dashboard was designed iteratively using data from contextual inquiry interviews to develop firstly paper based prototypes and then a visual design (using an interactive web based interface). Brief rapid feedback interviews were conducted with nurses to evaluate paper prototypes. After two rounds of feedback, a final prototype was identified that was then used to construct a web based interactive prototype (using a widely available web page prototyping tool www.invisionapp.com). The prototype included functionality to; enable the movement between different visualizations of the data (switching between a line graph and a bar graph), the ability to display data for different vital signs (using the radial buttons), and display of information in both the front screen of the patient EHR and in the patient notes display for vital signs (an example of the dashboard display for weight is in Figure 3). It included very basic decision support, by indicating to the user when the patient's weight gain or vital signs were outside either recommendations by guidelines or MD orders. The interactive prototype was evaluated for usability.

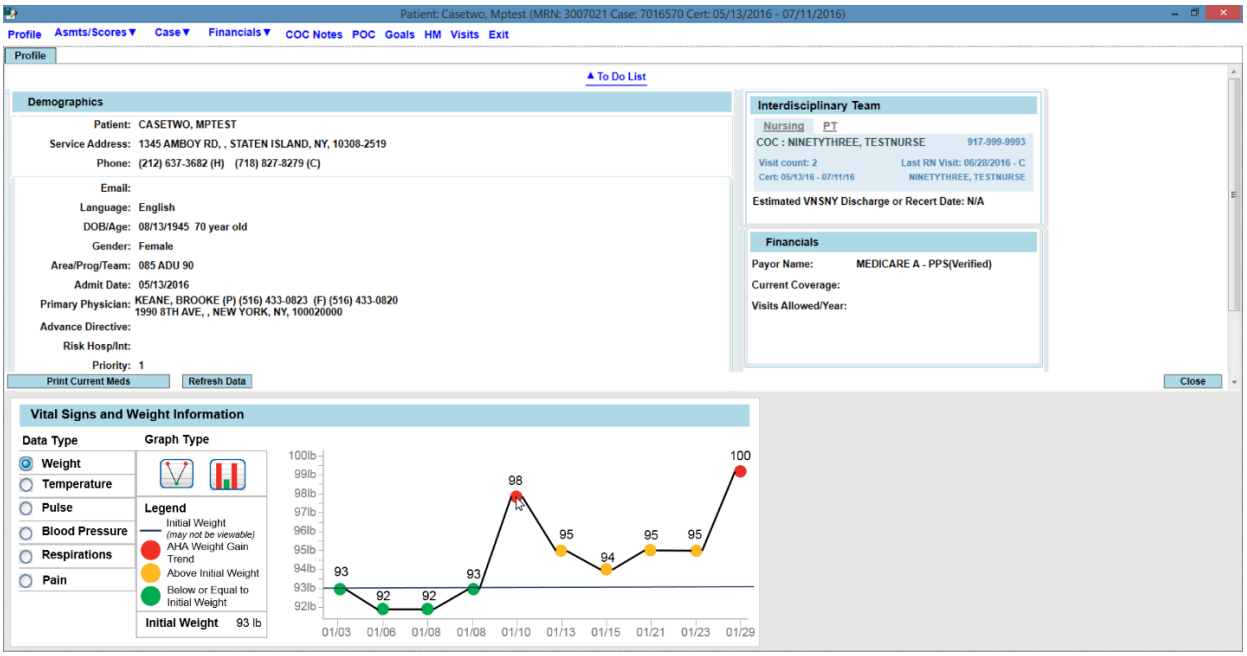


Figure 3: Example of prototype dashboard

Usefulness: The majority of nurses (91%) were able to start using the dashboard immediately and were able to easily switch between data elements using the radial buttons. They were also able to use the icons to navigate between the line and bar graphs (96%). Nurses liked the functions in the dashboard that enabled them to see trends for vital signs over time, without having to search back through previous notes. In addition, the color coding of the data, (red, yellow, green) were positively received, with nurses noting that this helped them to pay attention to the data. They also expressed a desire to see patient's blood glucose measures incorporated into the graphs. There was some disagreement on the overall usability of the system, with some nurses suggesting that they found it easy to use without training and others stating their need for training and practice in order to ensure they were comfortable to use functionality. User reactions seemed to vary depending on their level of computer literacy and dashboard functionalities.

Satisfying: SUS score: The dashboard had a mean score of 73.2 (SD 18.8) and a median score of 70. In general, it is considered that usable products should have SUS scores above 70, with a score of 73, the dashboard could be considered to have good usability.[40]

QUIS: Means of QUIS ratings of Dashboard Prototype (Scale from 1-9: word on the left = 1, word on the right =2) are summarized in Table 3. Elements of the dashboard given the highest ratings were terminology and dashboard information (mean 7.8; SD 1.1) and the dashboard screen (mean 7.7; SD 1.2). Elements with the lowest ratings were the overall user reactions (mean 6.1; SD 1.0) and learning the dashboard (mean 7.5; SD 1.3).

	N	Mean (SD)
System Usability Scale (SUS)	22	73.2 (18.8)
Questionnaire for User Interaction Satisfaction (QUIS)		
Overall User Reactions	***20***	***6.1 (1.0)***
Terrible/Wonderful	21	7.1 (1.5)
Frustrating/Satisfying	21	5.8 (2.6)
Dull/Stimulating	21	5.4 (2.6)
Difficult/Easy	21	6.1 (2.5)
Rigid/Flexible	20	6.5 (2.1)
Dashboard Screen	***22***	***7.7 (1.2)***
Characters on screen hard/easy to read	22	8.0 (1.4)
Highlighting helpful/unhelpful	22	8.0 (1.2)
Screen layouts helpful never/always	22	7.4 (1.3)
Sequence of screens confusing/clear	22	7.3 (1.7)
Terminology and Dashboard Information	***22***	***7.8 (1.5)***
The use of terminology inconsistent/consistent	22	7.8 (1.4)
Terminology relates to work never/always	22	7.7 (1.5)
Learning the Dashboard	***22***	***7.5 (1.3)***
Learning the dashboard difficult/easy	22	7.5 (1.3)
Exploration of features by trial and error discouraging/encouraging	22	7.5 (1.4)
Tasks performed in straightforward manner never/always	22	7.5 (1.4)
Dashboard Capabilities	***22***	***7.6 (1.4)***
Dashboard speed too slow/fast enough	22	8.0 (1.40
Ease of operation depends on level of experience never/always	22	7.1 (2.0)

Table 3: Summary of SUS and QUIS scores

Usability: Heuristic Evaluation**:** The dashboard had few usability issues in the categories of information coding, spatial organization, visibility of system status, orientation and aesthetic and minimalist design. The main issues were related to the flexibility and efficiency of use and user control and freedom. Overall, the usability issues identified by the experts were classified as being either a minor usability issue (n=5) or a cosmetic problem only (n=12). The specific issues raised included placing the icons for graph selection (line graph or bar graph) closer to the graph and using a contrasting color to highlight which graph is selected. They also made suggestions regarding the icon for the line graph, greying in the legend so it is clear that it is not interactive, and the ability to choose a) more data on one graph and b) the ability to show a broader or narrower date range.

Task Analysis: Nurses' average time to complete all dashboard tasks was 5.7 minutes (SD=2.4) compared to 1.4 minutes (SD=0.6) for the expert users. All nurses took more time than the average expert user. The average time deviation between nurses and

expert users was 4.3 minutes (SD=2.4). The ability of nurses to navigate the dashboard and extract accurate data from the displays was assessed in the task form. Questions specifically related to their ability to interpret data in the dashboards had accuracy ratings between 50% and 100%. Only 50% of users were able to accurately identify a patient's temperature, and 100% were able to identify their weight.

4.2 Discussion

The purpose of this pilot and feasibility study was to design dashboards that could provide feedback on quality metrics related to the care of patients with CHF to nurses working in a home care setting. The key aim of this project was to develop a dashboard that nurses would find useful, actionable and relevant to their practice. In addition, we also explored individual differences in comprehension of visualized data, which has implications for the future design of technology systems such as dashboards.

The results from our focus groups highlighted the importance that nurses attach to being able to see and visualize trended data on a patient's vital signs, across visits and with inbuilt alerts (decision support) to indicate when a patient's measures are outside those of existing clinical guidelines.[27] We also identified the need for such data to be readily and easily available to nurses during their clinical workflow; previous research using clinical dashboards to provide feedback to clinicians has highlighted the importance of not requiring users to actively search for the information to inform their practice.[41] In addition, the findings from the experimental online survey highlight significant individual differences in nurses' ability to comprehend visualized data, depending on both their numeracy and graph literacy. Of particular note are the significant issues with data comprehension exhibited by nurses who had both low numeracy and low graph literacy; our data highlighted that for these individuals presenting information in format such as a bar graph (where feasible) may have the highest likelihood of improving comprehension.

Given these findings, specific design elements of the prototype dashboard developed and evaluated in this study include; incorporation into the patient note when a nurse opens up their case (so that the data is included in the workflow and there is no additional work for the nurse to find this information), visualization of key data in 2 different formats (bar graph or line graph), the ability of the user to choose what information to view, and how they wish that data to be displayed, and some basic decision support. The decision support included a) an indication if the patient's weight had increased significantly enough to cause concern as indicated by the clinical guideline,[27] or b) if their blood pressure reading was outside of MD recommended parameters.

The usability evaluation of the prototype dashboard suggested that overall it was usable, nurses liked the functionality and rated it positively on both evaluation questionnaires and through feedback. However, before the dashboard could be implemented into a practice setting it would need to undergo some further modifications; in response to the heuristic evaluation this would include modifications to the interface to make it more explicit to the user which parts of the dashboard were interactive, and to

include further data on blood glucose measures. In addition, provision of training to nurses on the dashboard functionality and use would be beneficial.

The findings of our study highlight the importance of understanding user information needs, and individual variation in user characteristics, when designing technology such as dashboards. Existing research into the use of HIT systems indicates that often technology systems are designed without full consideration of the needs of clinicians; they often do not take user workflow into account and their introduction may lead to a number of unintended consequences.[42 43]

Study Limitations: This study was predominantly conduced in one home care agency, and was limited to nurses who work in a homecare setting. The results may not be generalizable to other health care settings or other populations of clinicians.

4.3 Conclusions

The use of dashboards to provide feedback to clinicians is increasing; our research highlights that it is important to question the assumption that using visualizations of data will automatically improve clinicians' ability to both understand that data and act upon it. It also highlights the importance of including users in the design process for technology solutions in clinical practice settings, alongside formal methods for evaluating usability of such systems, before they are widely introduced into practice. The dashboards in this study were designed to provide information to home care nurses that they consider to be important to enable them to work efficiently, in a format that maximizes nurses comprehension of the data (by enabling them to choose the format of data presentation).

4.4 Implications

The prototype dashboard designed in this study can be used as the basis for further research, to explore the effect of implementation on patient care processes and outcomes in a homecare setting. In addition, further research to explore the nature of individual differences in comprehension of data visualization displays, including with a wider population of clinicians, and using different types of data display, would be warranted.

5. LIST OF PUBLICATIONS AND PRODUCTS

Journal Papers

Dowding D, Merrill JA, Onorato N, Barrón Y, Rosati R, Russell D. The impact of home care nurses' numeracy and graph literacy on comprehension of visual display information: Implications for Dashboard Design. Journal of the American Medical Informatics Association 2017:Apr 27; doi: 10.1093/jamia/ocx042 [Epub ahead of print]

Dowding D, Russell D, Onorato N, Merrill JA. Perceptions of home care clinicians for improved care continuity: A focus group study. Journal for Healthcare Quality *(In press: accepted for publication May 18, 2017)*

Conference Oral Presentations

Accepted/Presented

Dowding D, Onorato N, Barrón Y, Merrill JA, Rosati RJ, Russell D. Supporting home care nurse decision making at the point of care through clinical dashboard design. H3IT. Home Healthcare and Hospice Information Technology Conference. Chicago, US. November 11, 2016.

Dowding D, Onorato N, Barrón Y, Merrill JA, Rosati RJ, Russell D. Using feedback intervention theory to guide clinical dashboard development: A case study in home care. ENRS, Philadelphia, PA. April 5-7 2017.

Dowding D, Onorato N, Rosati RJ, Barrón Y, Merrill J, Russell D. Using feedback intervention theory to design clinical dashboards. Medical Informatics Europe, Manchester UK, April 24-26, 2017.

Submitted and under review:

Dowding DW, Russell D, Jonas K, Onorato N, Barrón Y, Merrill JA, Rosati RJ. Does nurses' level of numeracy and graph literacy impact comprehension of quality targets? Findings from a survey of home care nurses. American Medical Association Annual Meeting, Washington DC, November 2017.

Posters

Accepted/Presented

Dowding D, Onorato N, Merrill J, Russell D. Improving quality of care for heart failure patients in a home care setting: what feedback information do nurses need? American Medical Association Annual Meeting, Chicago, November 12-16, 2016.

Russell D, Onorato N, Barrón-Vaya Y, Merrill J, Dowding, DW. The impact of nurses' graph literacy and numeracy on comprehension of visualized feedback information. American Medical Association Annual Meeting, Chicago, November 12-16, 2016.

Russell D, Onorato N, Merrill J, Barrón-Vaya Y, Dowding D. Nurses' graph literacy and its impact on comprehension of visualized feedback information. Academy Health Annual Research Meeting. Boston MA, June 26-28 2016.

Submitted and under review:

Dowding DW, Merrill JA. Heuristics for evaluation of dashboard visualizations. American Medical Association Annual Meeting, Washington DC, November 2017.

References

1. Ivers N, Jamtvedt G, Flottorp S, et al. Audit and feedback: effects on professional practice and healthcare outcomes. Cochrane Database of Systematic Reviews 2012;**6**:Art. No.: CD000259.
2. Harris-Kojetin L, Sengupta M, Park-Lee E, et al. Long-term care providers and services users in the United States: Data from the National Study of Long-Term Care Providers, 2013–2014. National Center for Health Statistics Vital Health Statistics 2016;**3**(38).
3. Jencks S, Williams M, Coleman E. Rehospitalizations among patients in the Medicare fee-for-service program. The New England Journal of Medicine 2009;**360**(14):1418-28.
4. Fortinsky R, Madigan E, Sheehan T, et al. Risk factors for hospitalization in a national sample of medicare home health care patients. Journal of Applied Gerontology 2014;**33**(4):474-93.
5. National Institute for Health and Clinical Excellence. Chronic heart failure. Management of chronic heart failure in adults in primary and secondary care. Manchester, UK: NICE, 2010.
6. Scottish Intercollegiate Guidelines Network. Management of chronic heart failure. A national clinical guideline. Edinburgh, UK, 2007.
7. Scott I. What are the most effective strategies for improving quality and safety of health care? Internal Medicine Journal 2009;**39**(6):389-400.
8. Daley K, Richardson J, James I, et al. Clinical dashboard: use in older adult mental health wards. The Psychiatrist 2013;**37**:85-88.
9. Bennett JW, Glasziou PP. Computerised reminders and feedback in medication management: a systematic review of randomised controlled trials. Medical Journal of Australia 2003;**178**:217-22.
10. Van Der Meulen M, Logie RH, Freer Y, et al. When a graph is poorer than 100 words: A comparison of computerised natural language generation, human generated descriptions and graphical displays in neonatal intensive care. Applied Cognitive Psychology 2010;**24**(1):77-89.
11. Hutchinson J, Alba JW, Eisenstein EM. Heuristics and biases in data-based decision making: Effects of experience, training, and graphical data displays. Journal of Marketing Research 2010;**47**(4):627-42.
12. Harvey N. Learning judgment and decision making from feedback. In: Dhami M, Schlottmann A, Waldmann M, eds. Judgement and Decision Making as a Skill Learning, Development and Evolution. Cambridge, UK: Cambridge University Press, 2013:199-226.
13. Gaissmaier W, Wegwarth O, Skopec D, et al. Numbers can be worth a thousand pictures: Individual differences in understanding graphical and numerical representations of health-related information. Health Psychology 2012;**31**(3):286-96.
14. Okan Y, Garcia-Retamero R, Cokely ET, et al. Individual Differences in Graph Literacy: Overcoming Denominator Neglect in Risk Comprehension. Journal of Behavioral Decision Making 2012;**25**(4):390-401.

15. Ziemkiewicz C, Ottley A, Crouser R, et al. Understanding visualization by understanding individual users. IEEE Computer Graphics and Applications 2012;**32**(6):88-94.
16. Berg Marc. *Rationalizing Medical Work*. Cambridge, Massachusetts: MIT press, 1997.
17. Dowding D, Mitchell N, Randell R, et al. Nurses' use of computerised clinical decision support systems: a case site analysis. Journal of Clinical Nursing 2009;**18**(8):1159-67.
18. Dowding D, Randell R, Mitchell N, et al. Experience and nurses use of computerised decision support systems. Studies in Health Technology & Informatics 2009;**146**:506-10.
19. Okan Y, Galesic M, Garcia-Retamero R. How People with Low and High Graph Literacy Process Health Graphs: Evidence from Eye-tracking. Journal of Behavioral Decision Making 2016;**29**(2-3):271-94.
20. Okan Y, Garcia-Retamero R, Galesic M, et al. When Higher Bars Are Not Larger Quantities: On Individual Differences in the Use of Spatial Information in Graph Comprehension. Spatial Cognition and Computation 2012;**12**(2-3):195-218.
21. Gaissmaier W, Wegwarth O, Skopec D, et al. Numbers can be worth a thousand pictures: individual differences in understanding graphical and numerical representations of health-related information. Health psychology : official journal of the Division of Health Psychology, American Psychological Association 2012;**31**(3):286-96.
22. Lopez KDPMPH, Wilkie DJPRNF, Yao YP, et al. Nurses' Numeracy and Graphical Literacy: Informing Studies of Clinical Decision Support Interfaces. Journal of Nursing Care Quality April/June 2016;**31**(2):124-30.
23. Boltz M, Capezuti E, Fulmer T, editors. *Evidence-Based Geriatric Nursing Protocols for Best Practice*. 4th Edition ed. New York, NY, USA: Springer Publishing Company, 2011.
24. Lindenfeld J, Albert NM, Boehmer JP, et al. Executive Summary: HFSA 2010 Comprehensive Heart Failure Practice Guideline. Journal of Cardiac Failure 2010;**16**:475-539.
25. National Heart Foundation of Australia and the Cardiac Society of Australia and New Zealand (Chronic Heart Failure Guidelines Expert Writing Panel). Guidelines for the prevention, detection and management of chronic heart failure in Australia.: National Heart Foundation of Australia, 2011.
26. Sobolewski S. The challenge of improving transitional care: lessons learned in a home healthcare agency. Home Healthc Nurse 2011;**29**(10):636-44.
27. Yancy CW, Jessup M, Bozkurt B, et al. ACCF/AHA guideline for the management of heart failure: a report of the American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines. Journal of the American College of Cardiology 2013;**62**:e147-239.
28. Burns Wendell J, Holtzblatt K, Wood S. *Rapid Contextual Design: A How to Guide to Key Techniques for User-Centered Design*. Burlington, MA: Morgan Kaufman, 2004.

29. Bednar P, Welch C. Contextual inquiry and requirements shaping. In: Wojtkowski W, Wojtkowsi G, Lang M, et al., eds. Information Systems Development: Challenges in practice, theory and education: Springer Verlag, 2009.
30. Lipkus IM, Samsa G, Rimer BK. General Performance on a Numeracy Scale among Highly Educated Samples. Medical Decision Making 2001;**21**(1):37-44.
31. Galesic M, Garcia-Retamero R. Graph literacy: a cross-cultural comparison. Medical decision making : an international journal of the Society for Medical Decision Making 2011;**31**(3):444-57.
32. Morae® v3.3 [program]. 3.3 version. Michigan, US: Techsmith Corporation.
33. QUIS: Questionnaire for User Interaction Satisfaction. Secondary QUIS: Questionnaire for User Interaction Satisfaction. http://www.lap.umd.edu/QUIS/.
34. Chin J, Diehl V, Norman K. Development of an instrument measuring user satisfaction of the human-computer interface. Proceedings of the SIGHCI conference on human factors in computing systems, 1988:213.
35. Harper B, Slaughter L, Norman K. Questionnaire administration via the WWW: A validation and reliability study for a user satisfaction questionnaire. WebNet 97, Association for the Advancement of Computing in Education. Toronto, Canada, 1997.
36. Hortman P, Thompson C. Evaluation of User Interface Satisfaction of a Clinical Outcomes Database. Computers, Informatics, Nursing 2005;**23**(6):301-07.
37. Johnson T, Zhang J, Tang Z, et al. Assessing informatics students' satisfaction with a web-based courseware system. International Journal of Medical Informatics 2004;**73**:181-87.
38. Moumame K, Idri A, Abran A. Usability evaluation of mobile applications using ISO 9241 and ISO 25062 standards. SpringerPlus 2016;**5**:548.
39. Zhang J, Walji M. TURF: Toward a unified framework of EHR usability. Journal of Biomedical Informatics 2011;**44**:1056-67.
40. Bangor A, Kortum P, Miller J. An Empirical Evaluation of the System Usablity Scale. International Journal of Human-Computer Interaction 2008;**24**(6):574-94.
41. Dowding D, Randell R, Gardner P, et al. Dashboards for improving patient care: Review of the literature. International Journal of Medical Informatics 2015;**84**(2):87-100.
42. Ash JS, Marc B, Coiera E. Some Unintended Consequences of Information Technology in Health Care: The Nature of Patient Care Information System-related Errors. Journal of the American Medical Informatics Association 2004;**11**:104–12.
43. Kim M, Coiera E, Magrabi F. Problems with health information technology and their effects on care delivery and patient outcomes: a systematic review. Journal of the American Medical Informatics Association 2017;**24**(2):246-60.