Health Informatics Journal

Volume 23 Number 4 December 2017

Health Informatics Journal

Article

Health Informatics Journal
2017, Vol. 23(4) 249–259

DOI: 10.1177/1460458216645149
journals.sagepub.com/home/jhi
SAGE

A novel approach to population-based risk stratification, comprising individualized lifestyle intervention in Danish general practice to prevent chronic diseases: Results from a feasibility study

Lars Bruun Larsen, Jens Soendergaard, Anders Halling, Trine Thilsing and Janus Laust Thomsen
University of Southern Denmark, Denmark

Abstract

Early detection of patients at risk seems to be effective for reducing the prevalence of lifestyle-related chronic diseases. We aim to test the feasibility of a novel intervention for early detection of lifestyle-related chronic diseases based on a population-based stratification using a combination of questionnaire and electronic patient record data. The intervention comprises four elements: (1) collection of information on lifestyle risk factors using a short 15-item questionnaire, (2) electronic transfer of questionnaire data to the general practitioners' electronic patient records, (3) identification of patients already diagnosed with a lifestyle-related chronic disease, and (4) risk estimation and stratification of apparently healthy patients using questionnaire and electronic patient record data on validated risk estimation models. We show that it is feasible to implement a novel intervention that identifies and stratifies patients for further examinations in general practice or behaviour change interventions at the municipal level without any additional workload for the general practitioner.

Keywords

electronic health records, health-care service innovation and information technology, information technology design and development methodologies, primary care, targeted prevention

Background

Interventions that contribute to a reduction in the increasing prevalence of lifestyle-related chronic diseases such as type 2 diabetes (T2DM), chronic obstructive pulmonary disease (COPD), and cardiovascular disease (CVD) are highly warranted.[1] Reducing the prevalence of chronic diseases calls for primary disease preventive efforts such as health-promotion activities and early detection

Corresponding author:
Lars Bruun Larsen, University of Southern Denmark, J.B. Winsløws Vej 9A, 5000 Odense, Denmark.
Email: lblarsen@health.sdu.dk

of people at risk for lifestyle-related chronic diseases.[2,3] Recent systematic reviews of general practice–based health checks suggest that people at high risk for a chronic disease may benefit from a targeted approach to health checks.[4,5]

To identify patients at high risk, the general practitioner (GP) needs systematically registered information on lifestyle risk factors in the electronic patient record (EPR).[3,6–11] However, GPs in Denmark and elsewhere lack systematically collected information on lifestyle.[12,13] This study aims to test the feasibility of a novel population-based risk stratification, comprising individual lifestyle intervention in primary care in Denmark for early detection of lifestyle-related chronic diseases using a combination of questionnaire and EPR data.

Methods

The intervention comprises four elements: (1) collection of information on lifestyle risk factors using a questionnaire with 15 validated items sent to patients listed with a GP, (2) electronic transfer of questionnaire data to the GPs' EPR, (3) identification of patients already diagnosed with a lifestyle-related chronic disease, and (4) risk estimation and stratification of apparently healthy persons using questionnaire and EPR data on validated risk estimation models (see Figure 1).

A total of 1200 individuals aged 39–59 years were randomly selected from the patient lists of four clinics in three regions in Denmark. No disease-related criteria for excluding a patient were defined prior to the study. Danish general practices are organized as publicly financed private clinics with a list system and an average of 1600 patients registered per GP and two GPs per clinic. Approximately half of the practices have one GP.[14] A total of 11 EPR systems for primary care are certified to comply with national standards for cross-sectional communication. All Danish GP clinics have fully implemented EPR systems in daily practice. The four participating GP practices consist of two or three GPs. One of the practices is situated in a provincial capital city with 250,000 inhabitants. The other three practices are in cities with populations of 3000 to 44,000 inhabitants. One practice is in an area mainly populated by persons with low socioeconomic status, another is in an area mainly populated by persons with high socioeconomic status, and the other two practices are located in areas with diverse populations. To increase the probability of reaching the aim of the feasibility study, the four clinics were chosen based on their former experience and engagement in research and quality development.

Collection of information on lifestyle risk factors using a questionnaire

The questionnaire was developed based on a literature search for validated questions used in the selected risk estimation models. Items on smoking habits, alcohol consumption, physical activity, weight, height, and family disposition for lifestyle-related disease (first-degree relatives >70 years of age) were drawn from the Danish National Health Survey.[15] Items on diet were drawn from the Swedish National Guidelines for Disease Prevention and translated directly from Swedish into Danish as the two languages are very similar.[16] Items on symptoms of COPD were drawn from a validated COPD population screener (PS),[7,16] translated from English to Danish using forward and back translation, and pilot tested among fellow research colleagues (see Appendix 1 for an English version of the questionnaire).

Questionnaire logistics

A letter of invitation and questionnaire were sent by the Danish Quality Unit of General Practice to participating patients by regular mail on behalf of each patient's GP. In the invitation letter, the

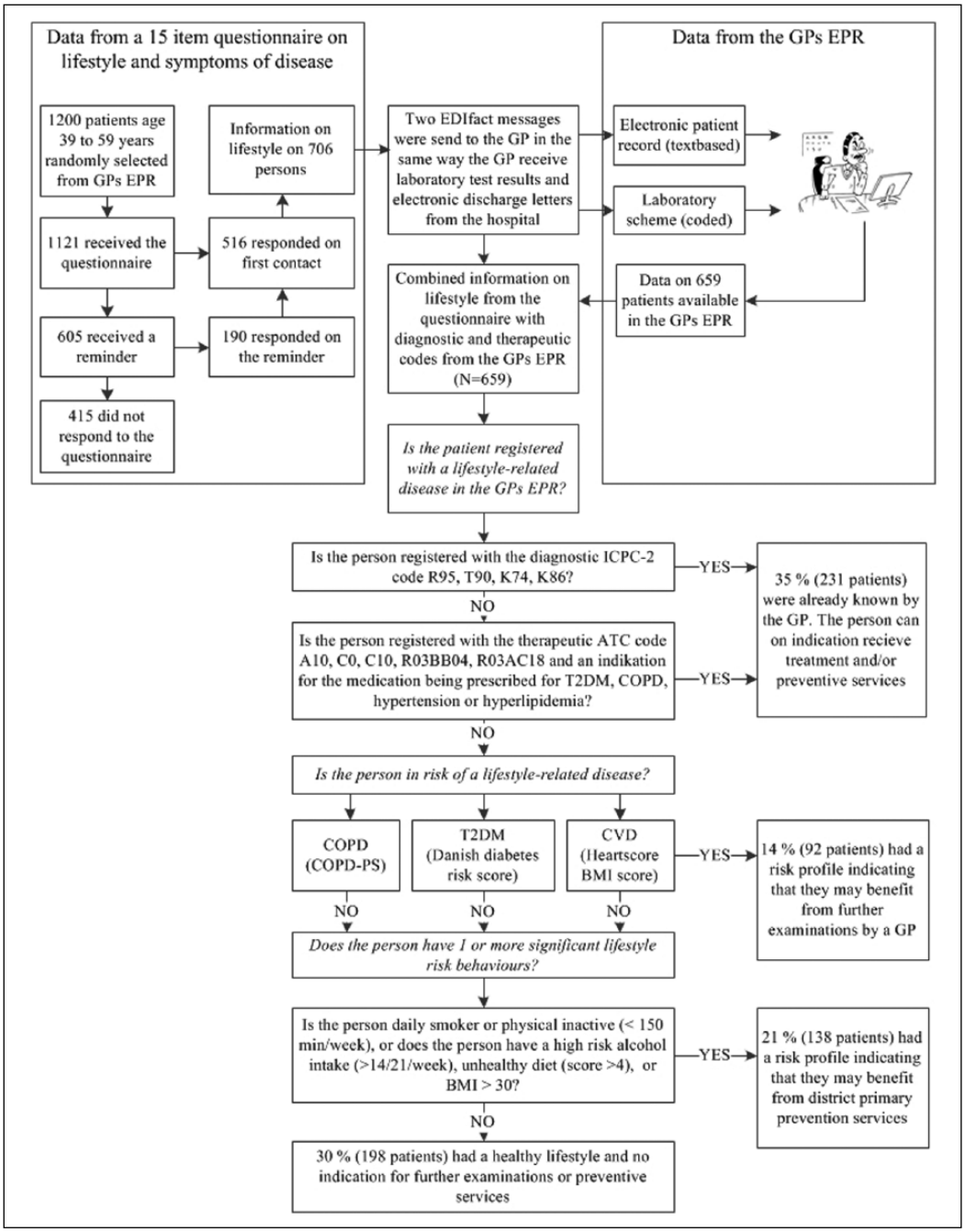

Figure 1. Flow diagram of the intervention.

patient was invited to fill in the questionnaire on paper and return it by mail or to fill it in electronically using the survey software solution SurveyXact (www.datafabrikken.dk). The invitation letter stated the purpose of the study, that it was voluntary to fill in and return the questionnaire, and that

the respondent was consenting to participate when returning the questionnaire. Non-response triggered a reminder after 3 weeks. The reminder was sent by regular mail. No further action was taken to increase the response rate. The participating GPs were consulted on the content and set-up of the invitation letter and questionnaire through mail correspondence and subsequent approval.

Transfer and synthesis of questionnaire and EPR data

The raw questionnaire data were processed to render them usable in a clinical context. Body mass index (BMI) was calculated based on self-reported weight and height (weight (kg)/height^2 (m)), and a dietary score and a score on physical activity were calculated following the recommendations in the Swedish National Guidelines for Disease Prevention.[16] The results from the questionnaire were subsequently transferred to the GPs' EPR systems as electronic data interchange (EDI) messages. For each patient, an EDI text message similar to a hospital discharge letter was transferred to the EPR, and an EDI coded message similar to biochemical laboratory results was transferred to the laboratory scheme. The coded message consisted of codes from the International Union of Pure and Applied Chemistry (IUPAC)/Nomenclature for Properties and Units (NPU) nomenclature and specific national codes for cross-sectional electronic communication within the Danish health-care system. The NPU terminology is a coding system and terminology for identification and communication of examination results from clinical laboratories in the health area and is supported by the International Federation of Clinical Chemistry (IFCC)-IUPAC (Sub)committee on NPU.

The coded EDI message contained information on smoking behaviour, BMI, units of alcohol consumption per week, minutes of physical activity per week, dietary score, and risk scores together with the result of the stratification stipulating whether the patient had been recommended to consult the GP or not.

Identification of patients already known to the GP

In 2005, the computer program Sentinel Data Capture was developed for Danish GPs with the purpose of collecting data for quality development and research.[17,18] Sentinel Data Capture is installed on a GP's PC or server and is designed to collect patient-specific data from the EPR. The collected data consist solely of structured data such as Anatomical Therapeutic Chemical (ATC) therapeutic codes on prescribed drugs, National Health Service disbursement codes, IUPAC/NPU laboratory codes, and International Classification of Primary Care (ICPC-2) encounters. Sentinel Data Capture was first used to collect key patient-specific data on diabetes care and seems to have contributed to improved care of patients with diabetes in general practice in Denmark.[18] The GPs have the possibility of using the whole range of ICPC-2 codes and are encouraged to assign ICPC-2 codes to all encounters.

In this study, patients already diagnosed with a lifestyle-related chronic disease or conditions such as hyperlipidaemia and hypertension were identified using a combination of ICPC-2 codes and therapeutic ATC codes for prescribed medicine, together with the indication for prescribing the medicine. Adding both the ATC therapeutic code and the indication for prescribing the medicine increased the likelihood that the medication was prescribed for, for example, COPD and not asthma. Patients with COPD were identified using the ICPC-2 code R95 or the ATC code category R03BB04 (tiotropium bromide) or R03AC18 (indacaterol) together with the indication text 'obstructive' or 'COPD' (see Table 2).

Patients identified by the validation algorithms were stratified to a group already known to be at risk for or have a lifestyle-related chronic disease and were as such assumed to be in some kind

Table 2. Validation algorithms.

Diagnosis	Diagnostic code(s)		ATC therapeutic code(s) for prescribed medicine and indicative text for the prescriptions[a]
Hypertension	K86, K87	*or*	C0 *BT*, *bt*, *Bt*, *ypert*, *ldot*, *LODTR*, *lotr*, *lodptr*, *bl. trykket*, *lodrtr*[b]
Hyperlipidaemia	T93	*or*	C10 *kolesterol*
COPD	R95	*or*	R03BB04 (tiotropium bromide), R03AC18 (indacaterol) *obstruktiv*, *KOL*
Type 2 diabetes	T90	*or*	A10 (drugs used for diabetes) *sukkersyge*, *diabetes*
CVD	K74, K76[c]		

ATC: Anatomical Therapeutic Chemical; COPD: chronic obstructive pulmonary disease; CVD: cardiovascular disease.
[a]The indicative text is in Danish and has not been translated as some indicative texts are only parts of the entire word and as such not translatable.
[b]The reason for the large number of indicative texts is misspelling by general practitioners (GPs) when indicating the purpose of the prescription.
[c]Diagnostic codes for ischaemic heart diseases are transferred to the GP's electronic patient record (EPR) system when the patients are discharged from the hospital following an angina episode or stroke. ATC codes for prescribed medicine will not provide further information.

of treatment, with either medication or behaviour change. These patients were not included in the subsequent estimation and stratification of people at risk for a lifestyle-related chronic disease.

Risk estimation and stratification of persons at risk

The subsequent stratification of persons at risk consisted of two steps. In the first step, patients at risk for a lifestyle-related chronic disease were identified using three validated risk scores: the COPD-PS screener, the Danish Diabetes Risk model, and the Heartscore BMI score.[7,9,19] All three models identify people at risk based on lifestyle risk factors and current EPR information only. It is important to keep in mind that risk estimation models such as the above are not intended to diagnose diseases. Diagnosis of disease calls for further examinations.

A COPD risk score was calculated based on the COPD-PS screener algorithm.[7] The COPD-PS screener uses an algorithm consisting of age, lifetime use of cigarettes, and symptoms from smoking to identify patients who may be offered spirometry to examine for COPD. The diabetes risk score was calculated based on the Danish Diabetes Risk score,[6] which uses an algorithm involving age, sex, BMI, known hypertension, physical activity, and parents having diabetes. The cut-off for being at risk for T2DM and COPD follows the recommendations in the Danish Diabetes Risk model and the COPD-PS screener, respectively. Patients with a COPD-PS score of 5 or above were considered eligible for spirometry. Patients with a diabetes risk score of 31 or above were considered eligible for HbA1c measurement.

Cardiovascular risk was calculated from the Heartscore BMI score,[9] which uses age, sex, smoking status, and BMI. The European Society of Cardiology recommends that this score be used to give a preliminary estimate of cardiovascular risk based on lifestyle risk factors alone, and as such, it is not a substitute for the full Heartscore. The Heartscore BMI score provides a figure for a 10-year CVD risk in the same way as the full Heartscore. In this study, the cut-off for being at risk

was established as a 5 per cent risk of dying from CVD within the next 10 years, which is the cut-off for determining whether the patient may benefit from further examinations and possibly also pharmacological treatment.

When one or more of the risk estimation models indicated a risk for a lifestyle-related chronic disease, the patient was stratified to further examinations by the GP to confirm or invalidate a possible diagnosis or provide further information for planning a treatment, whether behaviour change or pharmacological treatment. Patients stratified for further examinations by the GP were not included in the second step.

The second step identified patients with an unhealthy lifestyle with one or more risk factors. An unhealthy lifestyle was defined as daily smoking, alcohol intake of more than 14/21 (male/female) units of alcohol per week, an unhealthy diet (diet score <4 on a 12-point score drawn from the Swedish National Guidelines on Disease Prevention),[16] BMI >30, and/or physical activity <150 min/week. A patient with an unhealthy lifestyle with one or more risk factors was stratified to behaviour change interventions at the municipal level.

Finally, patients with no lifestyle-related chronic disease, no risk for a lifestyle-related chronic disease, and no unhealthy lifestyle risk factors were stratified to a group with a healthy lifestyle with no indication for further examination or intervention.

Ethics approval

The study was approved by the Danish Data Protection Agency. According to Danish regulations, approval from a health research ethics committee is not required for questionnaire surveys.

Results

Questionnaire response rates

A total of 1200 persons were selected for the study – 300 from each of the four clinics. Because the GPs receive a fee from the subsequent consultations, the Danish Consumer Ombudsman's view is that this kind of proactive contact with patients listed with the GP falls under the Danish Regulations on Marketing. We were therefore obliged to respect people who have actively chosen not to receive marketing by mail. Some 62 persons had actively chosen not to receive marketing by mail, and 17 persons had an unknown address. Hence, a total of 1121 persons received a questionnaire. Of these, 706 (63%) responded, 46 per cent at first contact and 17 per cent after one reminder. Of the 706 persons who returned the questionnaire, 62 per cent returned it in an enclosed envelope. The remaining 38 per cent filled in the questionnaire electronically.

The response rates of women and men were 67 and 59 per cent, respectively, and the response rate increased with age. The response rate in the four clinics ranged from 58 to 66 per cent. Some 685 persons filled in the entire questionnaire (see Table 3). The 21 persons who did not fill in the entire questionnaire were not included in the analysis.

Registration of lifestyle risk factors in GPs' EPRs

Among the respondents, the systematic registration of smoking and BMI in the GPs' EPR system was 17 and 23 per cent, respectively, prior to the intervention. Almost all respondents filled in the questionnaire, making it possible to increase the systematic registration of lifestyle risk factors in the EPRs up to 63 per cent (see Table 4).

Table 3. Results from the questionnaire.

% N = 685	Response rate	Self-rated health[a]	Daily smoker	Alcohol intake[b]	Unhealthy diet[c]	Physical inactivity[d]	BMI >30
Total	63	85	21	5	21	28	18
Male	59	82	22	3	31	25	18
Female	67	87	21	8	13	31	17
Age 39–44 years	57	87	17	3	24	26	14
Age 45–49 years	58	84	22	6	23	16	22
Age 50–54 years	69	87	23	7	20	20	16
Age 55–60 years	69	82	24	4	16	11	18
Clinic 1	58	82	26	5	25	26	29
Clinic 2	63	94	16	3	13	21	7
Clinic 3	64	79	25	7	25	27	20
Clinic 4	66	83	20	5	19	36	17
National Health Survey[e]		83	21	7	–	–	17

BMI: body mass index.
[a]Good, very good, or excellent self-rated health.
[b]>14/21 (men/women) units/week.
[c]Dietary score ⩽4 on a score from 0 to 12, where 12 represents a healthy diet.
[d]<150 min per week.
[e]Age interval from 45 to 54 years.[20]

Table 4. Number of responses on the individual risk factors.

Risk factor	Absolute number of responses (N = 706)
Smoking status	698
Alcohol consumption	700
Diet	705
Physical activity	705
BMI	700

BMI: Body mass index.

Stratification of respondents according to diagnosis and risk

It was possible to stratify 659 of the 706 respondents based on both the questionnaire and EPR data. The 24 persons whom we could not stratify because of missing EPR data are not included in the following analysis. Some 35 per cent of the respondents were already known by the GP as a patient with a diagnosis or receiving treatment for T2DM, CVD, or COPD and/or receiving treatment for hyperlipidaemia and/or hypertension. A total of 66 per cent received treatment for one condition, 21 per cent for two conditions, and 13 per cent for three or more conditions. Of the 35 per cent who were already known by the GP, 83 and 37 per cent received treatment for hypertension and hyperlipidaemia, respectively. A total of 14 per cent were identified as being treated for T2DM and 11 per cent for COPD, while 4 per cent were identified with ischaemic heart disease.

Some 14 per cent of the respondents (92 patients) had a risk profile indicating that they might benefit from further examinations by a GP, 16 per cent of whom (15 patients) were estimated to be

at risk for COPD. A total of 89 per cent (81 patients) were estimated to be at risk for T2DM and 24 per cent (22 patients) at risk for CVD. Some 74 per cent were estimated to be at risk by only one risk estimation model, 23 per cent by two models, and 3 per cent by all three models. A total of 21 per cent had a risk profile indicating that they might benefit from municipality primary prevention services but had no need for further examinations by the GP. Among these, 43 per cent were daily smokers, 8 per cent had a high-risk alcohol intake, 17 per cent had an unhealthy diet, 54 per cent were physically inactive, and 20 per cent had a BMI above 30. A total of 30 per cent (198 patients) had a healthy lifestyle and no indication for further examinations or preventive services.

Discussion

This study showed that it is feasible to stratify patients according to their risk of a lifestyle-related chronic disease without an additional workload for the GP and furthermore to target the services provided by the GP to one out of five 'healthy' patients.

Strengths and limitations

We managed to collect information on lifestyle from almost all respondents. We also managed to transfer information from the questionnaires to the GPs' EPR systems using the established system for electronic communication as both a text message and a coded message.

The four participating clinics were selected based on their former engagement in quality development projects. The clinics thus are most likely not representative of Danish general practice in terms of the use of ICPC-2 coding, prompting the question of whether the intervention is feasible in Danish general practices with less comprehensive ICPC-2 coding practices and less experience with information technology (IT)-based quality development tools. In this study, we used EDI messages and national communications standards such as the IUPAC/NPU nomenclature, which all Danish GPs can receive and send. The only active involvement from the GP was to approve the incoming EDI messages in the EPR. A meta-analysis has shown that the effectiveness of feedback to the GP depends on the baseline performance.[21] The lower the baseline recording, the greater the effect, suggesting that the impact of the intervention could be even larger in other less optimally performing Danish GP clinics than those participating in this study.

Patients already diagnosed with a chronic disease were identified based on ICPC-2 codes, ATC therapeutic codes, and the indication given for prescribing the drug. The strength of this way of validating diagnoses is high sensitivity because it captures patients receiving therapeutic treatment for a condition or a disease. The weakness is low specificity; patients are not identified by this model if they are either not registered with an ICPC-2 code but should have been, or do not receive therapeutic treatment for the condition or diagnosis. On the one hand, the registration of ICPC-2 codes among Danish GPs is currently not comprehensive, and ICPC-2 codes cannot in themselves identify all patients with a diagnosed disease. Close to 100 per cent of Danish GPs are enrolled in Sentinel Data Capture, but only 50 per cent of them register more than 70 per cent of their face-to-face consultations with an ICPC-2 code (May 2014). On the other hand, ATC therapeutic codes, together with an indication for prescribing the drug, will identify only patients receiving therapeutic treatment. Laboratory test results could add important information because they are currently used as diagnostic criteria for the diseases and conditions included in the study.[2,3,22] Algorithms using laboratory test results, ICPC codes, ATC therapeutic codes, and the indication for the prescription are thus warranted.

Even though we achieved a response rate of 63 per cent with a mail-distributed questionnaire, we lack information about lifestyle from 37 per cent of the patient population. Participation in

lifestyle surveys has declined during the last decade, and a response rate of 63 per cent with a low effort to increase the response rate shows that the intervention is perceived as valuable to the target population.[23] Two similar interventions in the Netherlands reached a response rate of 33 and 75 per cent, respectively.[24,25] According to the authors, the low response rate was primarily due to low confidence about online assessments among the respondents.[24]

Almost all Danes consult their GP within a time span of 3 years and around 80 per cent within one calendar year. Similar studies from New Zealand and Sweden have shown response rates of 91 and 70 per cent, respectively, using questionnaires distributed in the waiting room.[26,27] Questionnaires distributed in this way could, in combination with mail-distributed questionnaires, increase response rates and should be examined further.

Generalizability

The model is generalizable to other health-care systems, but it requires a primary care system with a highly developed IT infrastructure. This infrastructure should include ways to receive structured electronic information such as EDI messages from sources other than hospital-based laboratories and the possibility of combining this information with current EPR information to develop patient-specific risk profiles, either directly in the EPR or by a database. In the stratification of patients at risk for lifestyle-related chronic diseases, we focused our attention on lifestyle and lifestyle-based risk estimation models. Other questionnaires, risk estimation models, target groups, and focus areas such as cancer or mental illnesses can be used according to the specific purpose of the intervention.[8,26,27] However, laws and regulations on the privacy of health-related data may limit the possibilities of implementing a model like the one presented here.

Conclusion

It is feasible to implement an innovative intervention that identifies and stratifies patients for further examinations in general practice or behaviour change interventions at the municipal level without additional workload for the GPs. Using this model at the national level holds the potential to support GPs in preventing and treating lifestyle-related diseases and targeting diagnostic and lifestyle interventions towards high-risk patients. A large randomized intervention study is being planned that will combine the described model with targeted interventions in GP practice and the municipality.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

References

1. World Health Organization (WHO). *Global status on noncommunicable diseases*. Geneva: WHO, 2010.
2. Perk J, Backer GD, Gohlke H, et al. European Guidelines on cardiovascular disease prevention in clinical practice. *Eur Heart J* 2012; 33: 1635–1701.
3. Paulweber B, Valensi P, Lindström J, et al. A European evidence-based guideline for the prevention of type 2 diabetes. *Horm Metab Res* 2010; 42(Suppl. 1): S3–S36.
4. Si S, Moss JR, Sullivan TR, et al. Effectiveness of general practice-based health checks: a systematic review and meta-analysis. *Br J Gen Pract* 2014; 64(618): e47–e53.
5. Den Engelsen C, Koekkoek PS, Godefrooij MB, et al. Screening for increased cardiometabolic risk in primary care: a systematic review. *Br J Gen Pract* 2014; 64(627): e616–e626.
6. Glümer C, Carstensen B, Sandbaek A, et al. A Danish diabetes risk score for targeted screening: the Inter99 study. *Diabetes Care* 2004; 27(3): 727–733.

7. Martinez FJ, Raczek AE, Seifer FD, et al. Development and initial validation of a self-scored COPD Population Screener Questionnaire (COPD-PS). *COPD* 2008; 5(2): 85–95.
8. Alssema M, Newson RS, Bakker SJL, et al. One risk assessment tool for cardiovascular disease, type 2 diabetes, and chronic kidney disease. *Diabetes Care* 2012; 35(4): 741–748.
9. European Society of Cardiology. Heartscore BMI score, https://escol.escardio.org/heartscore/calc.aspx?model=europelow (accessed 5 December 2014).
10. Schuetz CA, Alperin P, Guda S, et al. A standardized vascular disease health check in Europe: a cost-effectiveness analysis. *PLoS ONE* 2013; 8(7): e66454.
11. Kahn R, Alperin P, Eddy D, et al. Age at initiation and frequency of screening to detect type 2 diabetes: a cost-effectiveness analysis. *Lancet* 2010; 375(9723): 1365–1374.
12. Mant J, Murphy M, Rose P, et al. The accuracy of general practitioner records of smoking and alcohol use: comparison with patient questionnaires. *J Public Health Med* 2000; 22(2): 198–201.
13. Rafter N, Wells S, Stewart A, et al. Gaps in primary care documentation of cardiovascular risk factors. *N Z Med J* 2008; 121(1269): 24–33.
14. Pedersen KM, Andersen JS and Søndergaard J. General practice and primary health care in Denmark. *J Am Board Fam Med* 2012; 25(Suppl. 1): S34–S38.
15. Sundhedsstyrelsen. *Den nationale sundhedsprofil 2010-Hvordan har du det?* Sundhedsstyrelsen, 2011, http://www.sst.dk/publ/Publ2010/CFF/Sundhedsprofiler/DenNationaleSHP.pdf (accessed 14 December 2012).
16. Svenska Socialstyrelsen. *Sjukdomsförebyggande metoder. Vetenskabeligt underlag för nationella riktlinjer*. Västerås: Svenska Socialstyrelsen, 2011.
17. DAK-E. Demonstration of sentinel data capture, http://demo.dak-it.dk/ (accessed 21 November 2012).
18. Schroll H, dePont Christensen R, Thomsen JL, et al. The Danish model for improvement of diabetes care in general practice: impact of automated collection and feedback of patient data. *Int J Fam Med* 2012; 2012: 208123.
19. Christensen JO, Sandbaek A, Lauritzen T, et al. Population-based stepwise screening for unrecognised Type 2 diabetes is ineffective in general practice despite reliable algorithms. *Diabetologia* 2004; 47(9): 1566–1573.
20. Sundhedsstyrelsen. *Danskernes Sundhed – Den Nationale Sundhedsprofil 2013*. København, 2014, https://sundhedsstyrelsen.dk/da/sundhed/planlaegning-og-beredskab/kommuner/~/media/8538E83A23B64880B3960909F85FED4D.ashx (accessed 9 November 2014).
21. Ivers N, Jamtvedt G, Flottorp S, et al. Audit and feedback: effects on professional practice and healthcare outcomes. *Cochrane Database Syst Rev* 2012; 6: CD000259.
22. Global Initiative for Chronic Obstructive Lung Disease. *Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease*, http://www.goldcopd.org (2014, accessed 16 September 2014).
23. Galea S and Tracy M. Participation rates in epidemiologic studies. *Ann Epidemiol* 2007; 17(9): 643–653.
24. Der Meer VV, Nielen MM, Drenthen AJ, et al. Cardiometabolic prevention consultation in the Netherlands: screening uptake and detection of cardiometabolic risk factors and diseases – a pilot study. *BMC Fam Pract* 2013; 14(1): 29.
25. Godefrooij MB, van de Kerkhof RM, Wouda PJ, et al. Identification of risk factors in primary care: development, implementation and yield of a cardiometabolic health check. *Fam Pract*. Epub ahead of print 9 September 2011. DOI: 10.1093/fampra/cmr068.
26. Blomstrand A, Ariai N, Baar A-C, et al. Implementation of a low-budget, lifestyle-improvement method in an ordinary primary healthcare setting: a stepwise intervention study. *BMJ Open* 2012; 2(4): e001154.
27. Goodyear-Smith F, Warren J, Bojic M, et al. eCHAT for lifestyle and mental health screening in primary care. *Ann Fam Med* 2013; 11(5): 460–466.

Appendix 1

Questionnaire

1. In general, would you say that your health is? (Excellent, Very good, Good, Fair, Poor)
2. What is your height in cm? (Height in cm)
3. What is your weight in kg? (Weight in kg)
4. Do you smoke? (Daily smoker, Occasional smoker, I quit smoking less than 6 months ago, I quit smoking more than 6 months ago, Never smoked)
5. During the past 4 weeks, how much of the time did you feel short of breath? (None of the time, A little of the time, Some of the time, Most of the time, All of the time)
6. Do you ever cough up any 'stuff', such as mucus or phlegm? (No, never; Only with occasional colds or chest infections; Yes, a few days a month; Yes, most days a week; Yes, every day)
7. Please select the answer that best describes you in the *past 12 months*. I do less than I used to because of my breathing problems. (Strongly disagree, Disagree, Unsure, Agree, Strongly agree)
8. How many units of alcohol do you consume in a regular week? (Number of units)
9. How often do you consume 4 or more units of alcohol (if female) or 5 units of alcohol (if male) on the same occasion? (Daily, Weekly, Monthly, Rarely, Never)
10. How often do you consume fruits (fresh, frozen, canned, or juice/smoothie)? (Twice a day or more often, Once daily, Some days during the week, Once a week or less often)
11. How often do you consume vegetables (fresh or frozen)? (Twice a day or more often, Once daily, Some days during the week, Once a week or less often)
12. How often do you consume fish and shellfish as the main dish? (Three times a week or more often, Two times a week, Once a week, A couple of times during 1 month or less often)
13. How often do you consume sweets, cakes, chocolate, or soda? (Daily, Almost every day, A couple of times during the week, Once a week or less often)
14. For how many hours during a week do you exercise (doing sports, running, bicycling, etc.)? (0 min (I do not exercise), Less than 30 min, 30–60 min (½ to 1 h), 60–120 min (1 to 2 h), More than 120 min (2 h or more))
15. For how many hours do you perform light exercise (walking, lawnmowing, etc.)? (0 min (I do not exercise), Less than 30 min, 30 to 60 min (½ to 1 h), 60 to 90 min (1 to 1½ h), 90 to 150 min (1½ to 2½ h), 150–300 min (2½ to 5 h), More than 300 min (5 h or more))

Health Informatics Journal

Article

Health Informatics Journal
2017, Vol. 23(4) 260–267

DOI: 10.1177/1460458216647089
journals.sagepub.com/home/jhi
SAGE

Under-coding of secondary conditions in coded hospital health data: Impact of co-existing conditions, death status and number of codes in a record

Mingkai Peng, Danielle A Southern, Tyler Williamson and Hude Quan
University of Calgary, Canada

Abstract

This study examined the coding validity of hypertension, diabetes, obesity and depression related to the presence of their co-existing conditions, death status and the number of diagnosis codes in hospital discharge abstract database. We randomly selected 4007 discharge abstract database records from four teaching hospitals in Alberta, Canada and reviewed their charts to extract 31 conditions listed in Charlson and Elixhauser comorbidity indices. Conditions associated with the four study conditions were identified through multivariable logistic regression. Coding validity (i.e. sensitivity, positive predictive value) of the four conditions was related to the presence of their associated conditions. Sensitivity increased with increasing number of diagnosis code. Impact of death on coding validity is minimal. Coding validity of conditions is closely related to its clinical importance and complexity of patients' case mix. We recommend mandatory coding of certain secondary diagnosis to meet the need of health research based on administrative health data.

Keywords

coding validity, hospital discharge data, secondary conditions (hypertension, diabetes, obesity and depression)

Introduction

Administrative health data including hospital discharge abstract database (DAD) have been widely collected and analyzed for various purposes, including disease surveillance, case-mix costing, tracking healthcare system performance, policy-making and research.[1,2] The Public Health Agency

Corresponding author:
Hude Quan, Department of Community Health Sciences, University of Calgary, 3280 Hospital Dr. NW, Calgary, AB T2N 4Z6, Canada.
Email: hquan@ucalgary.ca

of Canada (PHAC) has created the Canadian Chronic Disease Surveillance System (CCDSS) to conduct disease surveillance for 11 chronic conditions using administrative health data, such as physician claims data, hospital DAD.[3] CCDSS provides valuable information on national prevalence of chronic conditions and shows comparable trend with results from national survey. However, it should be cautious to interpret the results as underreporting of condition or misclassification of condition might underestimate or overestimate the disease prevalence. Under-coding of conditions, especially for asymptomatic conditions, has been identified as a major issue for administrative health data.[4] Hypertension, diabetes, obesity and depression were found to have no or protective effects on hospital death when administrative health data were used to conduct the risk adjustment for hospital mortality.[5] Use of administrative health data could result in underestimating prevalence for certain conditions, such as obesity.[6]

In Canada and many other countries, administrative hospital data are produced by health information professionals through review, abstraction and coding of data from inpatient charts following hospital discharge. According to Canadian coding standards, codes for the main diagnosis, any pre-admission or post-admission comorbidities and service transfer are mandatory while codes for secondary diagnoses not requiring clinical evaluation, therapeutic treatment, or increased nursing care and monitoring are optional.[7] It has been suggested that this could lead to incompleteness for asymptomatic conditions or conditions mainly treated in primary care settings.[8] Furthermore, coders are also subject to time constraints (usually 15–20 min for one medical chart) due to a high volume of work.[9] This could also impact the number of codes in each record.

In this study, we focused on four commonly under-coded secondary conditions (hypertension, diabetes, obesity and depression) in DAD. We hypothesized that when diagnosis information was transferred from chart to coded data, coding validity for asymptomatic conditions with modest clinical acuity in DAD could be impaired if its associated conditions are coded. Based on chart review data, we used logistic regression to identify any conditions listed in Charlson and Elixhauser comorbidities indices that are documented together with the four study conditions.[10,11] We examined the coding validity of the four study conditions related to whether their co-existing conditions were coded, whether the patient died in hospital and the total number of diagnosis codes recorded in a DAD record.

Methods

Data source

We randomly selected around 4000 records for patients aged $\geqslant$ 18 years and discharged between 1 January 2003 and 30 June 2003 from the four adult teaching hospitals in Alberta, Canada. There were at least 1000 records selected from each hospital, 4007 records in total. Each record was coded by the professionally trained health record coders using the Canadian coding standard for International Statistical Classification of Diseases and Related Health Problems (ICD), 10th Revision, Canada (ICD-10-CA). Canadian coding standard is maintained and developed by Canadian Institute of Health Information (CIHI) based on the ICD-10 developed by World Health Organization (WHO). Since 2001, Canada has used the ICD-10-CA coding standard for coding the DAD data. Minor amendments of coding standard were developed including wording or example changes and modification to reflect or clarify new directions. In DAD, there are 25 diagnosis code fields and 12 coding types. All diagnoses or conditions coded in the DAD must be assigned a diagnosis type.[7]

Based on the validated algorithms, we identified the conditions listed in Charlson and Elixhauser comorbidity indices.[12] The Charlson and Elixhauser comorbidities include 31 conditions and are two commonly used instruments for risk adjustment analyses.

Chart review

Two professionally trained reviewers reviewed all 4007 medical charts including a thorough review of the chart cover page, discharge summaries, narrative summaries, pathology reports (including autopsy reports), trauma and resuscitation records, admission notes, consultation reports, surgery/operative reports, anesthesia reports, physician daily progress notes, physician orders, diagnostic reports and transfer notes to check whether conditions listed in Charlson and Elixhauser comorbidities indices were documented. The process took approximately 1 hour for each chart. Detailed description about the process of chart review can be found in our previous publication.[4]

Statistical analysis

Identification of the co-existing conditions for the four conditions in the chart

Based on the chart review data, we developed logistic regression models via least absolute shrinkage and selection operator (LASSO) to identify any other conditions that were documented together with the four study conditions. The LASSO is a shrinkage and selection method for regression models that can be described as a constraint on the sum of the absolute values of the modeling parameters.[13] The LASSO allows for accurate estimation of model parameters and shrinks estimates of non-important parameter to zeros for automated variable selection. The independent variables used in the model are age, sex, and the remaining 30 conditions. Any conditions with nonzero estimates in the model were deemed as associated conditions.

Assessing validity

Sensitivity and positive predictive value (PPV) were used to assess the validity using the condition defined by chart review as the gold standard. Sensitivity indicates the probability of a condition being coded when a patient has the condition documented in chart; PPV indicates the probability of a condition being documented in the chart when a patient has the condition coded in the DAD.

Results

The four conditions of primary interest were under-coded in the DAD compared to chart review data (Table 1). All the co-existing conditions identified from chart data were clinically related to their corresponding conditions. Prevalence of the four conditions is three- or four-fold high if their co-existed conditions were coded. For all the four conditions, sensitivity was improved if their co-existed conditions were coded in the DAD. Coding of co-existing conditions in the DAD had negligible impact on PPV for the four conditions. Overall, diabetes and hypertension had high sensitivity while obesity and depression had low sensitivity in hospital DAD (Table 1).

Death was a severe outcome of hospitalization with average number of diagnosis code of 9.79 (vs 5.06 for the alive cases) in the dataset. There were 105 cases of death with 11 cases of missing status. Prevalence of the four study condition is high if the status of death was recorded. Status of death in the DAD has non-significant impact on the sensitivity and PPV for the four conditions (Table 2).

The total number of diagnosis code coded in the DAD ranged from 1 to 25 with a median number of 4 (interquartile range (IQR): 2–7). The sensitivity increased with an increase of the total number of diagnosis codes (Figure 1). Difference of sensitivity between records with 2 diagnoses and $\geqslant$8 diagnoses was 53 percent for hypertension, 35 percent for diabetes, 29 percent for obesity and 27 percent for depression. PPV was not related to the number of diagnosis codes in the DAD.

Table 1. Validity for the four study conditions with and without its co-existing conditions.

Presence of co-existing conditions	Number of cases	Prevalence, % (95% CI)		Sensitivity	PPV
		Chart review	DAD	(95% CI)	(95% CI)
Hypertension					
All observations	4007	22.1 (20.9, 23.5)	30.2 (28.8, 31.6)	68.3 (65.6, 70.9)	93.1 (91.3, 94.7)
CeVD	182	61.0 (53.5, 68.1)	68.1 (60.8, 74.8)	83.9 (76.2, 89.9)	93.7 (87.4, 97.4)
Diabetes	506	55.3 (50.9, 59.7)	66.8 (62.5, 70.9)	78.4 (73.6, 82.7)	94.6 (91.3, 97.0)
MI	336	55.7 (50.2, 61.0)	64.0 (58.6, 69.1)	81.9 (76.0, 86.8)	94.1 (89.7, 97.0)
Obesity	74	63.5 (51.5, 74.4)	64.9 (52.9, 75.6)	89.6 (77.3, 96.5)	91.5 (79.6, 97.6)
Renal failure	196	63.8 (56.6, 70.5)	72.4 (65.6, 78.6)	83.1 (75.9, 88.9)	94.4 (88.8, 97.7)
⩾1 co-existing conditions	1000	53.5 (50.4, 56.6)	63.7 (60.6, 66.7)	79.0 (75.6, 82.1)	94.0 (91.7, 95.9)
Diabetes					
All observations	4007	12.6 (11.6, 13.7)	14.6 (13.5, 15.7)	84.6 (81.4, 87.4)	97.6 (95.9, 98.8)
CHF	254	32.7 (26.9, 38.8)	35.4 (29.6, 41.7)	91.1 (83.2, 96.1)	98.8 (93.5, 100)
Hypertension	887	31.6 (28.5, 34.7)	32.9 (29.8, 36.1)	94.5 (91.3, 96.8)	98.6 (96.4, 99.6)
MI	336	26.2 (21.6, 31.2)	29.2 (24.4, 34.3)	89.8 (82.0, 95.0)	100 (95.9, 100)
Obesity	74	40.5 (29.3, 52.6)	41.9 (30.5, 53.9)	96.8 (83.3, 99.9)	100 (88.4, 100)
Renal failure	196	40.3 (33.4, 47.5)	43.4 (36.3, 50.6)	91.8 (83.8, 96.6)	98.7 (93.1, 100)
⩾1 co-existing	1184	29.6 (27.1, 32.3)	31.6 (28.9, 34.3)	92.2 (89.1, 94.7)	98.3 (96.3, 99.4)
Obesity					
All observations	4007	1.8 (1.5, 2.3)	8.3 (7.5, 9.2)	18.6 (14.6, 23.2)	83.8 (73.4, 91.3)
COPD	349	5.7 (3.5, 8.7)	11.7 (8.6, 15.6)	34.1 (20.1, 50.6)	70.0 (45.7, 88.1)
Diabetes	506	5.9 (4.0, 8.4)	20.8 (17.3, 24.5)	23.8 (16.0, 33.1)	83.3 (65.3, 94.4)
Hypertension	887	5.3 (3.9, 7.0)	14.8 (12.5, 17.3)	29.0 (21.4, 37.6)	80.9 (66.7, 90.9)
⩾1 co-existing	1293	4.5 (3.4, 5.8)	14.1 (12.2, 16.1)	25.8 (19.6, 32.8)	81.0 (68.6, 90.1)
Depression					
All observations	4007	5.8 (5.1, 6.6)	11.9 (10.9, 12.9)	44.9 (40.3, 49.5)	91.5 (87.1, 94.7)
Alcohol abuse	184	22.8 (17.0, 29.6)	42.4 (35.2, 49.9)	51.3 (39.7, 62.8)	95.2 (83.8, 99.4)
COPD	349	10.3 (7.3, 14.0)	17.2 (13.4, 21.6)	51.7 (38.4, 64.8)	86.1 (70.5, 95.3)
Drug abuse	113	27.4 (19.5, 36.6)	42.5 (33.2, 52.1)	60.4 (45.3, 74.2)	93.5 (78.6, 99.2)
Dementia	96	15.6 (9.0, 24.5)	30.2 (21.3, 40.4)	48.3 (29.4, 67.5)	93.3 (68.1, 99.8)
Fluid and electrolyte disorder	225	7.6 (4.5, 11.8)	19.6 (14.6, 25.3)	34.1 (20.5, 49.9)	88.2 (63.6, 98.5)
Hypothyroidism	149	13.4 (8.4, 20.0)	22.1 (15.8, 29.7)	60.6 (42.1, 77.1)	100 (83.2, 100)
Psychoses	73	13.7 (6.8, 23.8)	27.4 (17.6, 39.1)	40.0 (19.1, 63.9)	80.0 (44.4, 97.5)
⩾1 co-existing	956	13.8 (11.7, 16.2)	24.6 (21.9, 27.4)	51.5 (44.9, 58.0)	91.7 (85.6, 95.8)

PPV: positive predictive value; CI: confidence interval; DAD: hospital discharge abstract database; CeVD: cerebrovascular disease; MI: myocardial infarction; CHF: congestive heart failure; COPD: chronic pulmonary disease; ⩾1 related: at least one of the co-existed conditions coded in the DAD record.
The co-existed conditions for hypertension are CeVD, diabetes, MI, obesity and renal failure; for diabetes, they are CHF, hypertension, MI, obesity, renal failure; for obesity, they are COPD, diabetes and hypertension; for depression, they are alcohol abuse, COPD, drug abuse, dementia, fluid and electrolyte disorder, hypothyroidism and psychoses.

Discussion

Coding validity of conditions in the DAD was related to its clinical significance and complexity of patients' case mix. Hypertension, diabetes, obesity and depression are generally secondary

Table 2. Validity for the four study conditions related to the status of death.

Study conditions	Death	Prevalence, % (95% CI)		Sensitivity (95% CI)	PPV (95% CI)
		Chart review	DAD		
Hypertension	Yes	31.4 (22.7, 41.2)	46.7 (36.9, 56.7)	65.3 (50.4, 78.3)	97.0 (84.2, 99.9)
	No	21.8 (20.6, 23.2)	29.7 (28.3, 31.1)	68.4 (65.6, 71.1)	92.9 (91.0, 94.6)
Diabetes	Yes	25.7 (17.7, 35.2)	30.5 (21.9, 40.2)	81.2 (63.6, 92.8)	96.3 (81.0, 99.9)
	No	12.2 (11.2, 13.3)	14.1 (13.0, 15.2)	84.9 (81.6, 87.8)	97.9 (96.2, 99.0)
Obesity	Yes	3.8 (1.0, 9.5)	9.5 (4.7, 16.8)	30.0 (6.7, 65.2)	75.0 (19.4, 99.4)
	No	1.8 (1.4, 2.3)	8.3 (7.5, 9.2)	18.3 (14.2, 22.9)	84.3 (73.6, 91.9)
Depression	Yes	7.6 (3.3, 14.5)	17.1 (10.5, 25.7)	38.9 (17.3, 64.3)	87.5 (47.3, 99.7)
	No	5.8 (5.1, 6.6)	11.7 (10.7, 12.8)	45.1 (40.5, 49.8)	91.6 (87.1, 94.8)

CI; confidence interval; DAD: hospital discharge abstract database; PPV: positive predictive value.
There were 105 records with status of death and 3891 records without status of death.

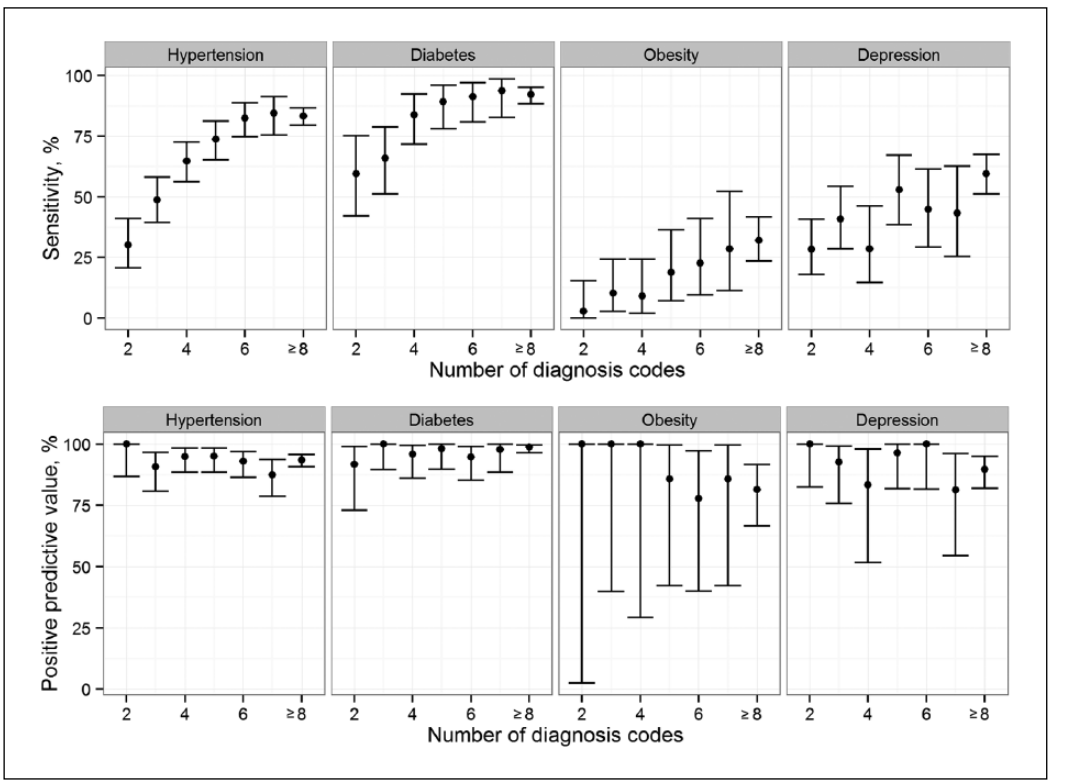

Figure 1. Sensitivity and positive predictive value of hypertension, diabetes, obesity and depression related to the number of diagnosis codes in hospital discharge abstract database.

diagnosis and their validity is affected by the coding of their co-existing conditions. The sensitivity for the four conditions increased as the total number of diagnosis codes in the record increased. Impact of death status on coding validity for the four conditions was minimal.

Coding validity is closely related to the clinical significance of a condition and its influence on length of stay, care received or therapeutic treatment during hospitalization. The four study conditions are generally secondary diagnosis during hospitalization, which are optional for coding according to the current coding standard in Canada.[7] Coding of co-existing conditions in the DAD was found to improve coding validity of the four study conditions. This provided a way to re-identify under-coded patients based on their comorbidities and improve the diseases surveillance. Lix et al.[14] found that inclusion of osteoporosis fracture and other fracture diagnosis in the administrative health data generally resulted in improved sensitivity of osteoporosis case-detection algorithm without loss of specificity.

Overall, hypertension and diabetes codes demonstrated better validity metrics than obesity and depression in the DAD. Hypertension and diabetes are the major risk factors for circulatory system diseases, which is the leading cause of hospitalization in Canada.[15] As a consequence, it seems more likely that hypertension and diabetes would be documented and coded in the DAD as shown in this study. Obesity was dramatically under-coded and underreported in the chart. The prevalence of obesity in chart review and inpatient DAD was 8.3 percent and 1.8 percent, respectively, much lower than 23.1 percent reported for the general population.[16] Furthermore, obesity had the lowest sensitivity among the four conditions. This is likely reflective of the fact that obesity generally fails to draw the physicians' attention or care on evaluation, treatment and management of main diseases during hospitalization.[17] Our previous study found that the higher the body mass index of patient, the more likely a diagnosis of obesity coded in the hospital DAD.[6] Depression was also under-coded in the DAD. It was noted that more than 90 percent of patients identified as having depression were receiving their care exclusively from a family physician. Under-coding of depression in DAD and poor documentation in medical chart could be related to the fact that treatment of depression during acute hospitalizations is suboptimal.[18]

The number of diagnosis codes in a record reflects the complexity of patient's case mix and quality of documentation of discharge summary. Data validity improved as the total number of diagnosis codes in DAD records increased. Increasing the number of diagnosis fields allowed in hospital data coding could enhance the completeness of coded clinical information in administrative health data. The WHO ICD, 11th version (ICD-11) topic advisory group on quality and safety recommended at least 15 secondary diagnosis fields to fully characterize clinical outcomes during hospitalization.[19] To fully describe a patient's health conditions, especially for chronic conditions, it might require more than one DAD record or the records collected over a specified time period (e.g. a few years). For example, it has been found that using a 1-year look-back period to identify comorbidity enabled better estimation of post-hospitalization mortality while using a look-back period longer than 1 year could help to accurately predict the readmission outcomes.[20]

Whether the patient died in hospital had minimal impact on coding validity for the study conditions. This is encouraging as administrative health data have been used to develop a series of indicators to calibrate the performance of hospitals and hospital mortality rate is one of the most important indicators.[21] To properly estimate this rate, it is required to conduct risk adjustment to account for the difference of patients' characteristics. This study provided evidence to support the use of administrative health data in development of health indicators related to mortality. However, it should be noticed that our study has a small number of records with patients died during their hospital stay.

The number of research studies based on administrative health data has been dramatically increasing in the recent years. Administrative health data have unique advantage, such as population coverage, low cost and timeliness. Administrative health data play a critical role in

community health assessment, disease surveillance, strategic planning, policy-making, service quality control and research. However, data validity remains questionable as the data collection priorities remain exclusively on billing or administrative purpose and not research. The current coding guidelines/practices hinder the completeness of inpatient data due to the focus on clinically significant reasons for the patient's admission or stay in hospital. Coding validity could be dramatically improved if all the conditions were coded regardless of whether those conditions are clinically implicated in the hospitalization. However, coding is a time-consuming and cost-intensive process. It is impossible to code all conditions, particular for complicated cases within the limited amount of time given to coders for each chart. It is suggested that some important chronic or modifiable conditions, such as hypertension and diabetes, should be coded as long as it was documented in the chart.

Limitations

This study has some limitations. First, we only examined the validity of four conditions with high prevalence. Other conditions having different clinical implication and resource use during hospitalization and prevalence might have different relationships between validity and existence of its associated comorbidities and patients' status of death. Second, we conducted our study based on the data from teaching hospitals. Teaching and nonteaching hospitalities vary in terms of severity and complexity of disease and case volume. Iezzoni et al.[22] reported that the validity of administrative health data varies between teaching and nonteaching hospitals. Third, we conducted our study based on a dataset from 2003. However, the fact that the dataset contained over 4000 records with their medical charts reviewed could be viewed as a strength. The coding guideline remains unchanged in the last 10 years. The process of chart review is costly and time-consuming as it provides more complete information on health records. So, despite these limitations, we believe that this study provides important insight into the data quality of the DAD and offers suggestions to potentially improve that data quality.

Conclusion

Coding validity of conditions is closely related to their clinical importance and the complexity of the patients' case mix. Hypertension, diabetes, obesity and depression are generally secondary diagnosis and optional for coding in hospital DAD according to the current coding standard. However, hypertension and diabetes, being common complications related to the leading cause of hospitalization in Canada, had better validity than obesity and depression. Furthermore, coding validity improved as the number of diagnosis codes in the record increased. We recommend the mandatory coding of certain secondary diagnosis to meet the increasing need of health service research conducted based on administrative health data. Use of hospital DAD only for surveillance faces the problem of underestimating the prevalence and incidence due to under-coding.

Funding

The author(s) received no financial support for the research, authorship and/or publication of this article.

References

1. Virnig BA and McBean M. Administrative data for public health surveillance and planning. *Annu Rev Public Health* 2001; 22: 213–230.
2. Robitaille C, Dai S, Waters C, et al. Diagnosed hypertension in Canada: incidence, prevalence and associated mortality. *CMAJ* 2012; 184: E49–E56.

3. Public Health Agency of Canada (PHAC). *The Canadian chronic disease surveillance system. Hypertension in Canada*, 2010. Ottawa, ON, Canada: PHAC, 2011.
4. Quan H, Bing L, Saunders LD, et al. Assessing validity of ICD-9-CM and ICD-10 administrative data in recording clinical conditions in unique dually coded database. *Health Serv Res* 2008; 43: 1424–1461.
5. Van Walraven C, Austin PC, Jennings A, et al. A modification of the Elixhauser comorbidity measures into a point system for hospital death using administrative data. *Med Care* 2009; 47: 626–633.
6. Martin BJ, Chen G, Graham M, et al. Coding of obesity in administrative hospital discharge abstract data: accuracy and impact for future research studies. *BMC Health Serv Res* 2014; 14: 70.
7. Canadian Institute of Health Information (CIHI). *Canadian coding standards for ICD-10-CA and CCI for 2015*. Ottawa, ON, Canada: CIHI, 2015.
8. Quan H, Parson G and Ghali W. Validity of information on comorbidity derived from ICD-9-CCM administrative data. *Med Care* 2002; 40: 675–685.
9. Lee DS, Donovan L, Austin PC, et al. Comparison of coding of heart failure and comorbidities in administrative and clinical data for use in outcomes research. *Med Care* 2005; 43: 182–188.
10. Charlson M, Szatrowski TP, Peterson J, et al. Validation of a combined comorbidity index. *J Clin Epidemiol* 1994; 47: 1245–1251.
11. Elixhauser A, Steiner C, Harris DR, et al. Comorbidity measures for use with administrative data. *Med Care* 1998; 36: 8–27.
12. Quan H, Sundararajan V, Halfon P, et al. Coding algorithms for defining comorbidities in ICD-9-CM and ICD-10 administrative data. *Med Care* 2005; 43: 1130–1139.
13. Hastie T, Tibshirani R and Friedman J. *The elements of statistical learning: prediction, inference and data mining*. 2nd ed. New York: Springer-Verlag, 2009.
14. Lix LM, Yogendran MS, Leslie WD, et al. Using multiple data features improved the validity of osteoporosis case ascertainment from administrative databases. *J Clin Epidemiol* 2008; 61: 1250–1260.
15. Public Health Agency of Canada. Leading causes of death and hospitalization in Canada, http://www.phac-aspc.gc.ca/publicat/lcd-pcd97/index-eng.php (accessed 25 March 2015).
16. Tremblay MS, Pérez CE, Ardern CI, et al. Obesity, overweight and ethnicity (Statistics Canada, catalogue 82–003). *Health Rep* 2005; 16: 23–34.
17. Kraschnewski JL, Sciamanna CN, Stuckey HL, et al. A silent response to the obesity epidemic: decline in US physician weight counseling. *Med Care* 2013; 5: 186–192.
18. Bilsker D, Goldner EM and Jones W. Health service patterns indicate potential benefit of supported self-management for depression in primary care. *Can J Psychiatry* 2007; 52: 86–95.
19. Drösler S, Romano PS, Sundararajan V, et al. How many diagnosis fields are needed to capture safety events in administrative data? Findings and recommendations from the WHO ICD-11 Topic Advisory Group on Quality and Safety. *Int J Qual Health Care* 2014; 26: 16–25.
20. Preen DB, Holman CD, Spilsbury K, et al. Length of comorbidity lookback period affected regression model performance of administrative health data. *J Clin Epidemiol* 2006; 59: 940–946.
21. Canadian Institute of Health Information (CIHI). *Health indicators 2012*. Ottawa, ON, Canada: CIHI, 2015.
22. Iezzoni LI, Shwartz M, Moskowitz MA, et al. Illness severity and costs of admissions at teaching and nonteaching hospitals. *JAMA* 1990; 264: 1426–1431.

Health Informatics Journal

Article

Health Information Exchange: What do patients want?

Health Informatics Journal
2017, Vol. 23(4) 268–278

DOI: 10.1177/1460458216647190
journals.sagepub.com/home/jhi
SAGE

Laura N Medford-Davis and Lawrence Chang
University of Pennsylvania, USA

Karin V Rhodes
Office of Population Health Management, Northwell Health/Hofstra Medical School, USA

Abstract

To determine whether emergency department patients want to share their medical records across health systems through Health Information Exchange and if so, whether they prefer to sign consent or share their records automatically, 982 adult patients presenting to an emergency department participated in a questionnaire-based interview. The majority (N = 906; 92.3%) were willing to share their data in a Health Information Exchange. Half (N = 490; 49.9%) reported routinely getting healthcare outside the system and 78.6 percent reported having records in other systems. Of those who were willing to share their data in a Health Information Exchange, 54.3 percent wanted to sign consent but 90 percent of those would waive consent in the case of an emergency. Privacy and security were primary concerns of patients not willing to participate in Health Information Exchange and preferring to sign consent. Improved privacy and security protections could increase participation, and findings support consideration of "break-the-glass" provider access to Health Information Exchange records in an emergent situation.

Keywords

electronic health records, emergency treatment, Health Information Exchange, health information technology, medical informatics

Introduction

A Health Information Exchange (HIE) is a secure repository of electronic health records (EHRs) organized by collaborative agreements between health systems, providers, and payers with the goal of providing access to critical elements of patients' medical records across multiple providers. The goals include expediting patient care, improving safety and quality, and care coordination. HIEs

Corresponding author:
Laura N Medford-Davis, Robert Wood Johnson Clinical Scholars Program, Department of Emergency Medicine, Perelman School of Medicine, University of Pennsylvania, 423 Guardian Drive, Blockley 1310, Philadelphia, PA 19104, USA.
Email: medford.davis@gmail.com

can also serve as a resource to track population health and public health. While HIEs vary in the information available, they typically allow providers to access a patient's treatment notes from other providers in other systems and integrate their own treatment plan with a patient's previous care plans in real time.

In emergency medicine, it is of particular importance to have easy access to a patient's prior health records as decisions about both testing and treatment need to be made in real time, usually without input from the patient's regular providers. While information that patients and family members may provide, if able, is vitally important, it is often incomplete.[1] It is notable that one-third of patients making a repeat visit to an emergency department (ED) also return to a different ED[2] where the records from their initial visit are not available. Access to information about current medications, recent electrocardiograms (EKGs), imaging results, and laboratory tests is hypothesized to improve the efficiency of an emergency evaluation, save time, and reduce costs, radiation exposure, and the potential for medical errors.[3] Without this access, EDs perform more tests to diagnose a patient's condition, which is not only expensive to the healthcare system but also increases a patient's exposure to harmful radiation.[4] Indeed, HIE has been shown to reduce repeat imaging and laboratory tests by 25 percent.[5,6]

Although much has been published about provider perspectives[6–8] and workflow integration issues,[9,10] less is known about patient preferences and concerns regarding the sharing of their personal health information across different health systems via an HIE. Prior surveys and focus groups have evaluated patients' opinions regarding sharing their records for research purposes,[11,12] but only a few studies have evaluated patient understanding of the risks and benefits of sharing their personal health records with their own healthcare providers for care coordination purposes through an HIE.[13,14] Knowledge of how patients perceive HIE can inform partnerships and data sharing agreements between HIEs and providers and payers which are required for successful HIE implementation.

The goal of this investigation was to survey ED patients who would be eligible to participate in an HIE being established in the area to determine (1) whether they currently have healthcare visits in multiple hospital systems in which their providers could benefit from information sharing, (2) whether they want their providers to share their health records across systems in the setting of an emergency visit, and (3) whether they prefer to sign consent or have their health records shared automatically.

Methods

Study design and population

Adult patients presenting to a large, urban, academic, tertiary hospital-based ED in the United States between the hours of 7 a.m. and midnight between 28 April and 11 August 2015 were approached by research assistants and verbally consented to participate in a brief questionnaire-based interview. This ED sees about 65,000 patients annually and is part of a three-hospital system with an extensive network of outpatient primary and specialty clinics. The patient population is primarily English speaking and predominantly insured through either commercial or public plans. ED physicians belong to a single salaried group. The system is in the process of converting to one electronic record system across all settings, but to date records can be shared for ED visits and discharge summaries between the three hospitals, but not for outpatient visits or the majority of inpatient notes.

Survey content and administration

The survey instrument (Supplementary Appendix 1) was designed to gather data on patients' needs and interest in care coordination, including preferred methods of contact, medication risks,

and opinions about HIE. The opinions about HIE form the basis for the data presented in this study. Questions were largely multiple choice, but two open-ended responses were included to preliminarily explore thought-processes when certain responses were chosen. The study and survey instrument were approved by the local institutional review board with a waiver of written consent.

Specific to the goals of this study, patients were asked questions to gauge whether an HIE would be relevant to them, including whether their usual provider used EHR, and whether all of their healthcare was in the same health system where the survey was conducted or whether they had also visited providers in other health systems. Patients were asked whether they expected their ED providers to be able to see their records from their other providers. Regarding HIE, patients were asked whether they would be willing to share their own health data in an HIE, and if so, which types of records they would want to share through an HIE. They were also asked, "Would you want your records to be shared automatically or would you prefer to sign permission before sharing your records?" A 10-percent sample of patients who said they preferred to sign permission were asked "In the case of a medical emergency where you were not able to give consent, would you want the doctors caring for you to be able to access your records automatically from other doctors and hospitals where you receive medical care?" to see whether they changed their answer from preferring to give consent to wanting their records shared automatically in the case of an emergency.

Demographic information was collected along with questions about general comfort with technology. Patients were asked whether they owned a text-capable cell phone, whether they were comfortable sending text messages, whether they had Internet in their home, and whether they sent emails. To minimize question fatigue and the burden on patient time, the patient's race and insurance were collected from the EHR. Based on the hypothesis that patients using controlled substances may wish to withhold that information from their providers, the survey also included the three-item Alcohol Use Disorders Identification Test (AUDIT-C).[15] If patients gave consent, their current medication record was also reviewed in the EHR to collect their actual use of controlled prescription medications (opiates and benzodiazepines).

After the authors piloted the survey with a small cohort of patients, questions were changed slightly to address common areas of confusion and create the final survey instrument used in this study. The choice "I don't know" or "unsure" was added as a possible response to whether the patient's primary physician used EHR. An initial third choice about whether patients would want to give consent once for HIE or would prefer to be re-consented each time their records were accessed through HIE was simplified to preference for either consent or no consent.

Patients who were under the age of 18 years, pregnant, or in police custody were not approached. Patients who were in mental or physical distress, with critical illness or injury, unable to provide informed consent due to intoxication, delirium, or other cognitive impairment, or non-English speaking were also excluded. Age was determined from the EHR. Police custody was visibly apparent from the presence of a police escort and/or handcuffs on the patient. Pregnancy and cognitive impairments were ascertained by asking the the patient's care provider. Eligibility was determined using an eligibility checklist. Consistent with Institutional Review Board (IRB) approval, verbal consent was obtained from all patients prior to the interview and patients were free to decline to participate, skip questions, or to stop their participation at any time. In order to comprehensively document participation rates, all patients presenting to the ED during the study period were entered in the study database, including those who declined to participate or met exclusion criteria.

The 12 research assistants are post-baccalaureate or undergraduate premedical students who are based in the ED and received training from two program coordinators on research ethics, obtaining informed consent, and survey delivery, as well as specific piloting of the survey instrument used in this study.

Data analysis

Chi-squared univariate tests compared willingness to share data in an HIE by patient demographic characteristics and other responses. A multivariable logistic regression analysis was then conducted to identify patient characteristics associated with willingness to share data in an HIE. We included variables with univariate p value <0.1 adjusted for all hypothetically relevant patient characteristics including age, race, educational level, use of technology, controlled substance use, and having healthcare providers outside of the health system in which the study was conducted (relevance of HIE). All analysis was performed using Stata (13.1, College Station, TX).

Results

Of 1152 potentially eligible ED patients approached, 1017 (88.3%) agreed to participate in the verbal questionnaire. Of these 1017, 982 (96.6%) responded to our primary outcome question of whether or not they were willing to share their data in an HIE and are therefore included in the analysis. Reasons recorded for not approaching patients (N = 321) included the patient was sleeping (N = 203; 63.2%), too ill or in too much pain (N = 105; 32.7%), under the influence of drugs/alcohol (N = 3; 0.9%), cognitively impaired (N = 5; 1.6%), or non-English speaking (N = 5; 1.6%). Of those who completed the survey (N = 982), 850 (86.6%) gave permission for medication review in their charts and 827 (84.2%) answered all questions included in the multivariable analysis.

Table 1[15,16] presents the demographics of the 982 participants, along with a univariate analysis of the association between these characteristics and their willingness to share data in an HIE. The median age of participants was 53 (interquartile range (IQR): 35.5–64). Of the patients who reported having outpatient providers in the same hospital system as the ED (N = 903), 883 (97.8%) expected their emergency physicians to be able to see their records from those outpatient physicians. While 492 of all 982 participants (50.1%) stated that all of their care was in the same hospital system as the ED, 770 (78.4%) reported having seen healthcare providers in other systems at least once.

The vast majority (N = 903; 92.0%) were willing to share their data in an HIE. Patients who were not willing were also less likely to have Internet at home, to use email, to have ever visited a provider outside of the hospital system, to not skip any questions, or to allow researchers to view their medication list in their EHR. Short qualitative reasons given by patients for not being willing to share their data in an HIE included concerns about privacy and confidentiality, security of their information, a belief that HIE would not benefit them because all of their care was already in one system, and a concern for not wanting to be locked out of insurance due to pre-existing conditions (Table 2).

Of those who wanted to share their health records through an HIE who also answered the question about consent (N = 897), 410 (45.7%) wanted their records shared automatically, whereas 487 (54.3%) wanted to sign consent before sharing their records. In all, 44 patients who were not willing to share their data in an HIE also answered the method of consent question, and three (6.8%) wished to share their records automatically if participating. Reasons patients gave for preferring to sign consent included privacy, awareness and control over who accesses their information, desire to keep some information secret from some doctors or hospitals, and desire to seek treatment only at one hospital system (Table 3). Patients willing to share their information automatically cited convenience, better quality of care when their providers had their information, and the possibility that they would be unable to sign in an emergency. Among the subset of patients who preferred to sign consent and were asked the follow-up question about emergency situations (N = 100), 90.0 percent said records should be shared automatically in the situation that they were unable to give consent.

Table 1. Univariate analysis of willingness to share health data in an HIE by patient demographics (N = 982).

	Total, N (%)[a]	Willing to share in an HIE, N (%)	Not willing to share in an HIE, N (%)	p value
Age (years)				0.738
18–49	396 (40.27)	366 (92.68)	29 (7.32)	
50–64	364 (37.10)	333 (91.53)	31 (8.47)	
65 and older	222 (22.63)	203 (91.03)	19 (8.97)	
Female	558 (57.06)	516 (92.47)	42 (7.53)	0.710
Male	420 (42.94)	385 (91.67)	35 (8.33)	
Race				0.276
Caucasian	362 (38.47)	338 (93.37)	24 (6.63)	
African American	538 (57.17)	488 (90.71)	50 (9.29)	
Hispanic	21 (2.23)	19 (90.48)	2 (9.52)	
Asian	20 (2.13)	20 (100.00)	0 (0.00)	
Education				0.217
Less than high school	111 (11.33)	100 (90.09)	11 (9.91)	
High school/GED	320 (32.65)	291 (90.94)	29 (9.06)	
Tech school	66 (6.73)	62 (93.94)	4 (6.06)	
Some college	202 (20.61)	181 (89.60)	21 (10.40)	
College degree+	281 (28.67)	266 (94.66)	15 (5.34)	
Owns a text message-capable cell phone	878 (89.05)	805 (91.52)	73 (8.31)	0.510
Uses text messaging	768 (78.13)	709 (92.32)	59 (7.68)	0.440
Has Internet at home	751 (76.17)	702 (93.48)	49 (6.52)	**0.001**
Uses email	716 (73.36)	667 (93.16)	49 (6.84)	**0.017**
Has a PCP	832 (84.47)	764 (91.83)	68 (8.17)	0.681
Usual source of care uses EHR	732 (77.96)	682 (93.17)	50 (6.83)	0.063
Unsure whether usual source of care uses EHR	167 (17.78)	147 (88.02)	20 (11.98)	
Has ever received healthcare in a different hospital system	770 (78.57)	719 (93.38)	51 (6.62)	**0.003**
Uses controlled prescription medications (opioids or benzodiazepines)	219 (26.58)	199 (90.87)	20 (9.13)	0.089
Did not give permission to view medication list	137 (14.02)	117 (85.40)	20 (14.60)	**0.002**
Risky alcohol use[b]	244 (24.90)	227 (93.03)	17 (6.97)	0.551
Skipped any questions	79 (8.04)	72 (7.59)	7 (21.21)	**0.005**

HIE: Health Information Exchange; PCP: primary care physician; EHR: electronic health record; GED: General Educational Development.

[a]Percent calculated from the total number who completed the combination of the two questions rather than all survey participants.

[b]Exceeds National Institute on Alcohol Abuse and Alcoholism (NIAAA)-sex/gender/age guidelines for safe use of alcohol.[15,16]

Note: bold values indicate statistical significance.

Almost all participants (N = 922; 94.0%), including 19 who were not willing to share their data in an HIE, wanted all possible records shared with their emergency physicians in an HIE, including physician notes, laboratory results, imaging results, hospitalization records, a medication list, a list of their medical problems, and a surgical history.

Table 2. Reasons for not wanting to share health data in an HIE (N = 76; 7.7%).

Theme	Examples
Privacy and confidentiality	"When information is shared it is no longer private." "Some things should be confidential." "Others besides physicians might snoop." "The only people that need to have my information are the people dealing with me." "If you decided to change your insurance company and they saw some disease you had in there they might not accept you because of that."
Security	"Information on the internet is not safe." "What if there was a breach of information?"
Number of doctors involved	"There would be too many disputes over the best method of care between physicians." "Each doctor should use their own judgment."
Applicability	"All of my doctors are at [this hospital]."

HIE: Health Information Exchange.

Table 3. Reasons for wanting to give consent to share health data in an HIE (N = 533; 56.2%).

Theme	Examples
Privacy, confidentiality, security	Personal privacy Seems safer, do not want to give it to all people
Knowledge and awareness	Know what is going on Would like to know when doctors are looking into something
Control over which physicians have records	Possibility of not liking physician or changing physician "In case I don't want that specific hospital to get the info." "Outside provider might just be a one-time thing." "It could be some information I don't want the doctor to know."

HIE: Health Information Exchange.

Results of an adjusted multivariable analysis are presented in Table 4. Patients who had never received care in another health system, who refused access to their electronic medication list, and who were uncertain whether their primary source of care used EHR remained significantly less likely to be willing to share their data in an HIE. Patient race, education level, adoption of technology for personal use, and abuse of alcohol or use of controlled prescription medications were not significant predictors of willingness to share health data in an HIE.

Discussion

This study found that the majority of patients are in favor of HIE, but about half want to control who accesses their information through explicit consent. The majority desiring consent are willing to waive it in the case of a true emergency however. Patients wary of HIE or wanting to maintain consent were primarily concerned about the privacy and security of their personal health information although a few other themes were elicited.

This study of patient perspectives on HIE is unique in its recruitment of a large number of patients actively seeking emergency care, whereas most prior surveys have queried potential healthcare consumers in non-healthcare settings[17–22] or in outpatient settings.[14] The ED in particular is a setting where patients see physicians they have never seen before who have no knowledge

Table 4. Multivariable model of factors associated with willingness to share health data in an HIE (N = 827).

Characteristic (comparison)	Odds ratio (95% confidence interval)
Age (10 years younger)	0.955 (0.782–1.166)
Female (male)	0.899 (0.498–1.621)
Race (Caucasian)[a]	
African American	0.942 (0.466–1.908)
Hispanic	0.623 (0.118–3.3.04)
Education (college degree or higher)	
Less than high school	0.937 (0.336–2.618)
High school/GED	1.122 (0.476–2.643)
Tech school	1.826 (0.372–8.967)
Some college	0.633 (0.273–1.467)
Does not own a text message-capable cell phone	1.548 (0.586–4.091)
Does not use text messaging	0.933 (0.410–2.120)
Does not have Internet at home	0.538 (0.260–1.116)
Does not use email	0.688 (0.318–1.490)
Does not have a PCP	1.267 (0.457–3.511)
Usual source of care does not use EHR (uses EHR)	0.761 (0.200–2.899)
Unsure whether usual source of care uses EHR	**0.464 (0.247–0.874)**
Has never received healthcare in a different hospital system	**0.450 (0.247–0.821)**
Uses controlled prescription medications (no controlled medications listed in EHR)	0.631 (0.326–1.224)
Did not give permission to view medication list	**0.366 (0.180–0.748)**
Risky alcohol use[b]	1.247 (0.615–2.530)

HIE: Health Information Exchange; PCP: primary care physician; EHR: electronic health record; GED: General Educational Development.

[a]Survey responses from 17 Asian participants were not included in logistic regression model due to perfect prediction of outcome (all 17 Asian participants were willing to share their data in an HIE).

[b]Exceeds NIAAA-sex/gender/age guidelines for safe use of alcohol.[15,16]

Note: bold values indicate statistical significance.

of their prior health history, giving HIE an especially high utility in the ED environment. The vast majority of ED patients surveyed were in favor of an HIE, higher than seen in prior studies.[13,14,21–23] The approximately 8 percent who were not willing to share their data in an HIE were less familiar with technology and might be less comfortable with the concept in general. In addition, patients who did not give permission for research assistants to view their medication list were also less likely to want their medical records shared through HIE. This remained significant in multivariable analysis, but it is unclear whether this question tested general distrust of allowing others to view health records, or evasiveness specific to hiding use of controlled substances. Those who skipped questions on this survey were also less likely to want to share their records through an HIE. Unfortunately, only four participants who skipped questions had responses to all variables in the multivariable analysis, all of whom were not willing to participate in HIE, so we were unable to test this association more definitively, but it suggests that general distrust or desire to keep personal matters private was a primary factor.

Patient comments from Tables 2 and 3 highlight distrust about sharing personal health information and electronic data security. Many patients feared that staff members besides their providers or insurers would abuse HIE to access their records inappropriately. Several prior studies have

shown that security and privacy are major barriers for patients,[3] and a study that gathered multiple stakeholder input in Taiwan found that patients have greater concerns about privacy than do physicians.[24] This study shows that this disconnect also needs to be addressed in the United States. One patient was also fearful of an insurer denying them coverage based on the HIE information, indicating that either patients are not aware that the Affordable Care Act outlawed denial of coverage for preexisting conditions, or that they do not trust its ability to enforce that rule. Patients also do not seem to recognize how much information their insurers already receive about their diagnoses and medications in order to process their claims.

In our study, patients were somewhat more likely (56% vs 46%) to say they preferred to provide consent before their records were shared via HIE as opposed to having them shared automatically. This is similar to prior surveys in other settings that have found 35–69 percent of participants wanting to sign consent for HIE participation.[14,21] However, among the subset of patients who preferred to give consent that were asked a follow-up question, 90 percent indicated that they would want their ED doctors to automatically access their health records in a medical emergency if they were unable to give consent and no one else could be reached. While respondents might have felt pressured to change their response to please researchers when asked a second time, our results are consistent with a prior study where support for access to records without consent increased from 35 to 93 percent when participants were presented with a scenario in which they were in an emergency situation and unable to provide their own medical information,[21] and another study where 90 percent supported access when the question was phrased as a medical emergency.[17] This finding highlights the tension between the Healthcare Insurance Portability and Accountability Act (HIPAA), HIE, and emergency situations.

It is the current standard of practice that patients in emergency situations who are unable to provide informed consent are presumed to consent to life-saving treatment. Therefore, it could be argued that "break-the-glass" HIE access should be a part of any emergency care where patients are unable to sign consent for access to their medical records. Further public education is needed about the current situation in some HIEs where opting out makes any medical records invisible, even in the case of true emergency scenarios such as unconsciousness after car accidents or medical catastrophes such as heart attack or stroke where access to medical records could be life-saving.

Of the patients surveyed in this urban ED, four out of five said they had providers and hence medical records in another health system, including 61.6 percent of patients who also reported that all of their care was in the current hospital system. These results suggest that a highly functional and accessible HIE would be of benefit to a majority of ED patients. Consistent with patient comments that loyalty to the current hospital system obviated their need to participate in HIE, patients who reported that they had never seen providers outside of the hospital system were less likely to be willing to share their data in an HIE, a characteristic that remained significant on multivariable analysis. Public education on the likelihood of being taken by ambulance to an ED in a different health system in the case of a true emergency could also improve support for HIE access in EDs.

Most patients (97.8%) expected their emergency physicians could already see records from their outpatient physicians in the same system, but in the health system where this study was conducted, emergency physicians can only see occasional notes about phone conversations but not full physician visit notes due to a difference in EHRs between settings and EHR uptake by outpatient providers, highlighting a disconnect between public expectations and the actual state of integrated EHR in the United States where only 54 percent of office-based physicians have achieved meaningful use of EHR,[22] and many different EHRs may be utilized within the same health system without seamless information sharing.[25]

Limitations

Although this study has a large sample size, it was performed at a single center and the results may not be generalizable to other areas of the country or to other ED populations. Nineteen percent (N = 190) of surveys had missing responses related to key variables and were therefore excluded from the final multivariable analysis, so there may be systematic differences between patients who answered all questions and those who did not. This was a qualitative study with both multiple-choice and open-ended responses, and not all open-ended responses were captured verbatim nor were they systematically coded. Future studies might consider a more in-depth discussion with the patients who are not willing to share their data in an HIE about their concerns, with coding of the exact responses to elicit themes. The study investigated patients' reported willingness to share their data in an HIE, but did not actually ask them to enroll. People's behavior in a real situation does not always match what they say they would do hypothetically. Although concerns regarding privacy were elicited from those not willing to share their data in an HIE and questions were asked regarding alcohol and controlled substance use, substance abuse, psychiatric history, and human immunodeficiency virus (HIV) or sexually transmitted infection history were not explicitly included in the choices of types of information to be shared through an HIE. Future research may wish to determine whether concerns about this type of more sensitive information increase overall privacy and security concerns. Nonetheless, results are strongly supportive of the acceptability of HIE to the ED population, particularly in a truly emergent situation.

Conclusion

Study results show that the majority of ED patients are in favor of HIE, and that consent to access HIE records should be explored for inclusion in the routine "permission to treat" forms that all emergency patients sign when seeking care in an ED. If patients opt-out, there should be an option that also gives patients the ability to make an exception (e.g. opt-in to HIE) in the case of life-threatening emergencies. The qualitative data support a need for patient education regarding the utility of HIEs and the role of insurers in HIEs, and a need for improvements in the security of HIEs and regulation on who can access HIEs.

Acknowledgements

The authors would like to thank Briah Paciotti, Sheila Kelly, Christopher Gibson, Kathryn Wanner, Cjloe Vinoya, Danielle Kovalsky, Felicia Wong, Anjerie Yohn, and the Emergency Medicine Academic Associate Program who helped conduct the surveys and enter the data. Research data and analyses are available upon request by contacting the authors.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The study was supported by a University of Pennsylvania grant (Rhodes, PI) from the Leonard Davis Institute of Health Economics, the Center for Clinical Epidemiology and Biostatistics, and the Center for Public Health Initiatives.

References

1. Roane TE, Patel V, Hardin H, et al. Discrepancies identified with the use of prescription claims and diagnostic billing data following a comprehensive medication review. *J Manag Care Pharm* 2014; 20: 165–173.
2. Duseja R, Bardach NS, Lin GA, et al. Revisit rates and associated costs after an emergency department encounter: a multistate analysis. *Ann Intern Med* 2015; 162: 750–756.
3. Rudin RS, Motala A, Goldzweig CL, et al. Usage and effect of health information exchange: a systematic review. *Ann Intern Med* 2014; 161: 803–811.
4. Lammers EJ, Adler-Milstein J and Kocher KE. Does health information exchange reduce redundant imaging? Evidence from emergency departments. *Med Care* 2014; 52: 227–234.
5. Yaraghi N. *The benefits of health information exchange platforms: measuring the returns on a half a billion dollar investment*. Brookings: Center for Technology Innovation, 2015.
6. Kierkegaard P, Kaushal R and Vest JR. How could health information exchange better meet the needs of care practitioners? *Appl Clin Inform* 2014; 5: 861–877.
7. Rudin RS. Why clinicians use or don't use health information exchange. *J Am Med Inform Assoc* 2011; 18: 529.
8. Thorn SA, Carter MA and Bailey JE. Emergency physicians' perspectives on their use of health information exchange. *Ann Emerg Med* 2014; 63: 329–337.
9. Unertl KM, Johnson KB and Lorenzi NM. Health information exchange technology on the front lines of healthcare: workflow factors and patterns of use. *J Am Med Inform Assoc* 2012; 19: 392–400.
10. Dullabh P and Hovey L. Large scale health information exchange: implementation experiences from five states. *Stud Health Technol Inform* 2013; 192: 613–617.
11. Weitzman E, Kelemen S, Kaci L, et al. Willingness to share personal health record data for care improvement and public health: a survey of experienced personal health record uses. *BMC Med Inform Decis Mak* 2012; 12: 39.
12. Kim K, Joseph J and Ohno-Machado L. Comparison of consumers' views on electronic data sharing for healthcare and research. *J Am Med Inform Assoc* 2015; 22: 821–830 .
13. Patel VN, Dhopeshwarkar RV, Edwards A, et al. Low-income, ethnically diverse consumers' perspective on health information exchange and personal health records. *Inform Health Soc Care* 2011; 36: 233–252.
14. Patel VN, Dhopeshwarkar RV, Edwards A, et al. Consumer support for health information exchange and personal health records: a regional health information organization survey. *J Med Syst* 2012; 36: 1043–1052.
15. Bush K, Kivlahan DR, McDonell MB, et al. The AUDIT alcohol consumption questions (AUDIT-C): an effective brief screening test for problem drinking. Ambulatory Care Quality Improvement Project (ACQUIP). Alcohol Use Disorders Identification Test. *Arch Intern Med* 1998; 158: 1789–1795.
16. US Department of Health and Human Services, National Institutes of Health, and National Institute on Alcohol Abuse and Alcoholism. *Helping patients who drink too much: a clinician's guide*. Bethesda, MD: National Institutes of Health, 2007.
17. Ancker JS, Edwards AM, Miller MC, et al. Consumer perceptions of electronic health information exchange. *Am J Prev Med* 2012; 43: 76–80.
18. Ancker JS, Silver M, Miller MC, et al. Consumer experience with and attitudes toward health information technology: a nationwide survey. *J Am Med Inform Assoc*. 2013; 20: 152–156.
19. Dimitropoulos L, Patel V, Scheffler SA, et al. Public attitudes toward health information exchange: perceived benefits and concerns. *Am J Manag Care* 2011; 17: SP111–SP116.
20. Wen KY, Kreps G, Zhu F, et al. Consumers' perceptions about and use of the Internet for personal health records and health information exchange: analysis of the 2007 Health Information National Trends Survey. *J Med Internet Res* 2010; 12: e73.
21. Dhopeshwarkar RV, Kern LM, O'Donnell HC, et al. Health care consumers' preferences around health information exchange. *Ann Fam Med* 2012; 10: 428–434.
22. O'Donnell HC, Patel V, Kern LM, et al. Healthcare consumers' attitudes towards physician and personal use of health information exchange. *J Gen Intern Med* 2011; 26: 1019–1026.

23. Simon SR, Evans JS, Benjamin A, et al. Patients' attitudes toward electronic health information exchange: qualitative study. *J Med Internet Res* 2009; 11: e30.
24. Wang JY, Ho HY, Chen JD, et al. Attitudes toward inter-hospital electronic patient record exchange: discrepancies among physicians, medical record staff, and patients. *BMC Health Serv Res* 2015; 15: 264.
25. Carter JH. What is the electronic health record. In: Carter JH (ed.) *Electronic health records: a guide for clinicians and administrators*. Philadelphia, PA: ACP Press, 2008, pp. 3–19.

Article

Health Informatics Journal

Health Informatics Journal
2017, Vol. 23(4) 279–290

DOI: 10.1177/1460458216647759
journals.sagepub.com/home/jhi
SAGE

Diagnosis of adverse events after hysterectomy with postoperative self-care web applications: A pilot study

Donna T Gilmour, Norman J MacDonald and Steven Dukeshire
Dalhousie University, Canada

Barbara Whynot
IWK Health Centre, Canada

Barry Sanders
University of British Columbia, Canada

John Thiel
University of Saskatchewan, Canada

Sony Singh and Craig Campbell
University of Ottawa, Canada

Krisztina Bajzak
Memorial University, Canada

Gordon Flowerdew
Dalhousie University, Canada

Abstract

Increased pressures from multiple sources are leading to earlier patient discharge following surgery. Our objective was to test the feasibility of self-care web applications to inform women if, when, and where to seek help for symptoms after hysterectomy. We asked 31 women recovering at home after hysterectomy

Corresponding author:
Norman J MacDonald, Dalhousie University, 5850/5890 University Avenue, Halifax, NS B3K 6R8, Canada.
Email: normmacdonald@gmail.com

at two centers to sign into a website on a schedule. For each session, the website informed them about normal postoperative symptoms and prompted them to complete an interactive symptom questionnaire that provided detailed information on flagged responses. We interviewed eight women who experienced an adverse event. Six of these women had used the web application regularly, each indicating they used the information to guide them in seeking care for their complications. These data support that self-care applications may empower patients to manage their own care and present to appropriate health care providers and venues when they experience abnormal symptoms.

Keywords
education of patients, hysterectomy, patient safety, postoperative complications, web-based applications

Introduction

Hysterectomy is one of the most common surgical procedures performed on women in developed countries.[1] There are 650,000–700,000 hysterectomies performed in Canada and the United States each year.[2,3] Hysterectomy can lead to major peri-operative morbidity, including urinary tract and bowel injury, infection, hemorrhage, thromboembolism, and death.[4] A prospective clinical audit of 1330 hysterectomies with a form specifically designed to record complications from hysterectomy demonstrated that 26.4, 23.9, and 17.0 percent of patients with abdominal, laparoscopic, and vaginal hysterectomies, respectively, experienced adverse events.[5] A larger prospective cohort study with detailed questionnaires on 5279 hysterectomies demonstrated complication rates of 19.2, 15.4, and 11.7 percent in patients with abdominal, laparoscopic, and vaginal hysterectomies, respectively.[6] Because adverse events following hysterectomy can be common and some of the adverse events can be serious, we need excellent pre-surgical/in hospital patient education by health professionals and reliable methods of providing accurate information to women and their families once they return home.

Adverse events and their consequences after elective hysterectomy can be significant and costly for patients, surgeons, the health care system, and society. Diagnosing and managing serious adverse events often requires multiple outpatient and inpatient laboratory, imaging, medical, and surgical resources.[3] These investigations and procedures are uncomfortable, associated with their own risks and side effects, and costly.[3] Major adverse events after hysterectomy can negatively and permanently impact an individual's length and quality of life. Even non-serious adverse events, such as urinary tract, wound, or pelvic infections, can cause discomfort, stress, and anxiety for patients. These less severe adverse events may also be inefficiently and ineffectively managed if patients present to inappropriate health care venues for evaluation and treatment. A recent retrospective chart review of laparoscopic hysterectomy and same-day discharge found that the most common reasons for visiting the emergency department within 48 h were pain and nausea.[7] The study concluded that detailed postoperative instructions may have prevented many of these early postoperative visits.[7]

Gap in recovery information provided to patients

At present, informing and educating patients about their surgery and postoperative recovery consists of preoperative visits with their surgeons and nurses along with verbal and written instructions provided at discharge. At most Canadian hospitals, preoperative teaching is done by a registered nurse usually during a pre-admission visit a week or more before scheduled surgery. The majority of postoperative teaching in hospital is provided by nurses and is usually done as the patient is

being discharged from hospital. Although this is the standard of care for Canadian postoperative patients, it is sometimes ineffective as women are often experiencing pain, and may be under the influence of anesthetic and narcotics, as well as feeling anxious or rushed about their discharge. Increased emphasis on cost-savings has reduced the amount of time spent in preoperative sessions with nurses.[8] About one-third of 102 hysterectomy patients surveyed felt that they had been provided insufficient information.[9]

Patients are increasingly supplementing verbal and written information with online resources. The use of Internet-based self-management tools has improved the health status of patients with chronic disease when compared to standard non-Internet self-management programs.[10] Online tools have been developed to educate women about family planning[11] and to help quit smoking during pregnancy.[12] Online patient portals developed by hospitals to manage chronic diseases such as diabetes in children have been studied and found to have potential.[13] Electronic support tools available for aiding in recovery from surgery include the website Hystersisters.com, which has information that is not created by health care professionals but is instead centered on peer-support. In a study of 137 women who used the Hystersisters.com website, women found the information about recovery helpful; however, 39 percent found some aspect (often negative postings from some members) of the website not helpful in their recovery.[14] One study demonstrated that patients who used Hystersisters.com valued some information more if it was provided by a member perceived as knowledgeable.[15] A meta-analysis of 191 studies found that increased patient education shortens length of stay in hospital by an average of 11.5 percent and has beneficial effects for recovery, pain, and psychological distress.[16] A recent randomized controlled trial has shown that orthopedic patients who were given supplemental online information from authoritative sources to research on their own prior to their surgery modified their decision on spinal versus general anesthesia.[17] All of these studies suggest there may be a gap in the knowledge that patients have about their surgery and that it may be possible to help fill that gap with authoritative and reliable online or mobile app resources. With adverse events after hysterectomy occurring at home being relatively common and potentially serious, we also need to ensure these resources provide accurate, timely, and reassuring information tailored to their concerning symptoms and type of surgery.

Efficient navigation of a complex health care system

The Canadian health care system—with its many types of care providers and multiple access points for care—may leave patients uncertain as to where they should go when they experience symptoms of adverse events at home after surgery.[18] Early presentation, diagnosis, and treatment of evolving adverse events are critical steps in limiting their short-term and long-term sequelae. Early diagnosis and treatment of an adverse event also has a positive effect on the health care system, as early intervention can often prevent a hospital admission. Currently, no formal screening mechanisms are in place to survey patients quickly, routinely, and comprehensively for specific symptoms of adverse events after hysterectomy and to advise them when and where to go for appropriate care.

Proposal to fill information gaps

In response to all of these information needs and health systems gaps, we created the SAFER (Studying Adverse Events From Elective Surgery Research) web application.[19] The application consisted of surgery-specific, interactive applications designed to inform and empower patients to better care for themselves after surgery. We conducted a small study on 31 abdominal hysterectomy patients to see if they would access the self-care application from home after surgery and use the information to assist them in caring for themselves. Here, we report on 11 of the 31

women who suffered adverse events during the study and how using the application impacted their care and outcomes.

Materials and methods

Gynecologic surgeons met and determined the initial approach and content for the abdominal hysterectomy applications. The content was broken down to include a 2-min checkup to screen for symptoms, recovery advice on what is normal and not normal (specific to the patient's day of recovery), as well as background educational content on hysterectomies. A medical writer and computer programmer transformed this information into the SAFER Abdominal Hysterectomy Self-care Web Application. Designed to be accessed regularly by patients recovering at home after surgery, the modules were concise and specific to abdominal hysterectomy and to the day of recovery after surgery. The self-care application was intended to be accessed daily during the first week at home and then twice a week for three more weeks. During each session, patients received brief, timely advice about how to care for themselves and were asked to answer 18 screening questions about specific symptoms after surgery (Appendix 1). Patients who gave a positive reply to a symptom question were provided additional information about the possible significance of the symptom and where they should go for further evaluation. They also had the option of accessing the symptom information without providing a positive reply to the screening question. Detailed information on the development, design, and implementation of the tool, including screenshots and patient feedback on the overall design, has been previously reported.[19] Here, we focus on the experience and outcomes of those who used the application and experienced an adverse event. These results will be used to support efforts to launch a larger study into the effect of these types of tools on adverse event rates and patient satisfaction.

After approval from the IWK Health Centre institutional review board, five women having abdominal hysterectomy enrolled in a small feasibility cohort. Women who had Internet access from home, who were comfortable using the Internet for basic services (e.g. banking, bill payment, and/or shopping), and who could speak and read English were included. These women were approached to join the study by a preoperative clinic nurse who was already scheduled to conduct the patient's regularly scheduled preoperative appointment. The total number of patients approached by the preoperative clinic nurses was not recorded.

A research nurse gave a brief introduction and demonstration on how to use the website to the five women prior to their surgery. While recovering at home after surgery, they were asked to sign-in to the SAFER modules and complete the specified activities. These activities consisted of reading the daily recovery advice, answering the 18 symptom questions (Appendix 1), and completing a short free-text entry if they had to seek health care advice or treatment for a problem. Other activities included a preoperative questionnaire to record their demographics, information sources, expectations, and a postoperative questionnaire to record their satisfaction and health care access.

At the conclusion of their participation, all five women provided detailed feedback by taking part in a semi-structured telephone interview. We took into account patient feedback from this small cohort of five patients by modifying the content and utility of the application based on their comments and suggestions. This detailed patient feedback from these five women has been previously reported by the authors.[19]

After obtaining institutional review board approval from the IWK Health Centre in Halifax and the Ottawa Hospital Research Institute in Ottawa, we tested the revised version of the online self-care application in a second small pilot cohort involving 26 women having abdominal hysterectomies at either of these two sites. These 26 women were asked to complete the same activities as the women in the feasibility cohort. For all women in the first (feasibility) cohort (n = 5) and the second

(pilot) cohort (n = 26), we collected baseline demographic and usage rates for the web application. Adverse event occurrence was determined by medical research staff completing a medical chart review and contacting each participant to capture adverse events not documented in their medical chart. All 31 women were contacted 6 weeks post-surgery to inquire if they had experienced any other problems that required them to access the health care system. Qualitative details about these health problems were recorded. With the sample sizes being too small to draw definitive comparisons between groups and the fact that the pilot cohort largely followed the same methodology with minor revisions following the feasibility stage, all 31 women's experiences are reviewed as one cohort in this article.

Results

Demographics

Figure 1 contains a flow chart of participant enrollment, compliance, and adverse event occurrence. Table 1 provides the demographic characteristics, broken down by adverse event occurrence, of the women who used the web applications and completed the demographic questionnaire.

Participants experiencing adverse events

Eleven (35%) of the 31 women who used the website suffered an adverse event after surgery. These adverse events consisted of one ureteral fistula, four wound complications, two vaginal

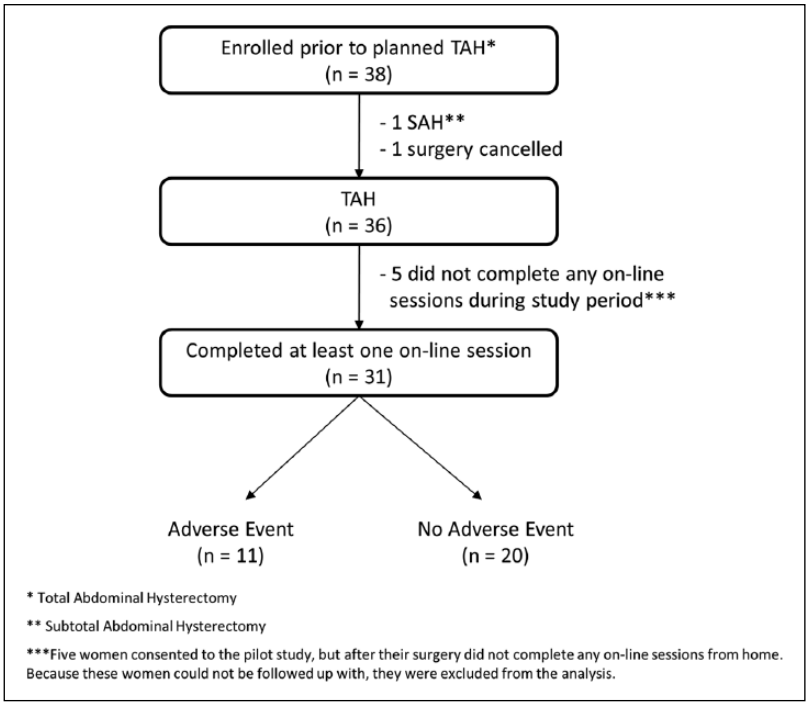

Figure 1. Flow chart of enrollment of study participants.

Table 1. Demographic information of women who used the postoperative self-care web applications and completed the demographic questionnaire.

Characteristic	Adverse event	No adverse event
	No. (%)	No. (%)
	(n = 9)[a]	(n = 18)[a]
Age (years)		
Under 40	1 (11.1)	6 (33.3)
40–49	5 (55.6)	8 (44.4)
50–59	3 (33.3)	4 (22.2)
Any previous surgical operation		
Yes	5 (55.6)	15 (83.3)
No	4 (44.4)	3 (16.7)
Education level		
Less than grade 12	0	1 (5.6)
Completed high school	0	1 (5.6)
Some college/university	0	3 (16.7)
Completed college/university	9 (100)	13 (72.2)
Annual household income		
Less than US$20,000	0	1 (5.6)
US$20,000–US$39,999	2 (22.2)	0
US$40,000–US$59,999	0	5 (27.8)
US$60,000–US$79,999	3 (33.3)	3 (16.7)
Over US$80,000	4 (44.4)	9 (50.0)
Distance >20 km from		
Hospital of surgery	4 (44.4)	10 (55.6)
Nearest hospital	3 (33.3)	5 (27.8)
Family doctor's office	4 (44.4)	6 (33.3)

[a]Two subjects from each group did not complete the demographic questionnaire.

vault complications (vault hematoma and vault cellulitis), and four urinary tract infections. No patient suffered more than one adverse event.

We were able to contact and interview 8 of the 11 women who experienced an adverse event. Three of the 11 women could not be reached for an in-depth interview. These three women had a urinary tract infection as their adverse event. For the eight women interviewed, Table 2 presents their diagnosed adverse event, the frequency of their web application usage, and whether the online material had influenced their decision to seek additional care.

Two participants who experienced an adverse event, rows 1 and 2 in Table 2, felt that the website did not affect their treatment decision, due to the symptoms being obvious or occurring at a regularly scheduled follow-up appointment with their own physician. The other six women who experienced an adverse event used the applications regularly and to guide their decision making about whether to seek additional care and where to go for this care. Two participants with an adverse event, rows 3 and 4 in Table 2, felt the website confirmed and reassured their own decision to seek help, with the second woman stating it hastened her decision. One participant experienced a ureteral fistula, row 5. After following the website information and presenting to an Emergency

Table 2. Number of online sessions completed by eight participants[a] who experienced an adverse event and whether the information influenced their decision to seek additional care.

	Adverse event	Completed online sessions (of 13)	Did the information influence your decision to seek additional care?
1	Wound hematoma	2	No
2	Superficial wound dehiscence	5	No
3	Wound infection	8	Yes
4	Wound hematoma	13	Yes
5	Ureteral fistula	11	Yes
6	Urinary tract infection	12	Yes
7	Vault hematoma	12	Yes
8	Vault cellulitis	13	Yes

[a]Of the 11 women who suffered an adverse event, 3 women with a urinary tract infection could not be reached for interview.

Department, she had an initial misdiagnosis of acute cholecystitis. She felt her uterovaginal fistula might have been diagnosed and treated sooner if the website had provided even more information on what her symptoms meant so that she could have communicated this to her health care providers. Other comments from these eight women include that the web application was helpful, easy to use, clear, and reassuring. They also noted the website was credible due to it coming from their health care centers and providers.

Barriers to use

Six of the 31 women (19%) did not use the web applications regularly (completed <50% of the sessions). To understand why women did not use the applications, we approached all six women for interviews and four agreed. Two of the women interviewed had also suffered an adverse event (Table 2, first two entries). For three of the women, their reasons for not using the online applications were related to issues of access. Two women reported that, because their computers were set up on a different floor than where they spent the bulk of their recovery time, it was difficult and/or painful to get to their computer. One woman experienced such debilitating pain (from her wound hematoma) that she could not get to her computer to complete the applications. The fourth woman reported being too overwhelmed by the demands of single parenting and postoperative recovery to be able to use the applications consistently.

Patient overall experience

The website's post-surgery questionnaire recorded patients' opinions about their experience during recovery, their satisfaction with their surgery outcome, and usefulness of the web applications. Women who experienced an adverse event tended to give higher scores for the questionnaire items assessing the helpfulness and usefulness of the web applications (Table 3). Fisher's exact test was performed on a comparison of Agreement ($\geqslant$4) versus No Agreement (<4) in the two groups. No significant differences were found as all p values were above 0.19 on this small sample size.

Table 3. Satisfaction with the applications for women who suffered and did not suffer an adverse event (AE), on a 5-point scale ranging from 1 (strongly disagree) to 5 (strongly agree).

Item	AE mean score (% 4 or 5)	No AE mean score (% 4 or 5)
	(n = 8)	(n = 16)
The SAFER website was useful in helping me decide whether or not to contact a health professional for symptoms I had	4.5 (75%)	3.9 (67%)
I used the SAFER website more than any other information source (not including health professionals) during my recovery	4.4 (88%)	3.9 (79%)
Overall, the SAFER website helped me in my recovery	4.4 (75%)	4.4 (83%)
The 2-min checkup[a] helped reduce my worry	4.3 (75%)	3.8 (67%)
The 2-min checkup[a] was useful in helping me to decide whether or not I should call a health professional for symptoms I had	4.5 (88%)	4.1 (75%)

SAFER: Studying Adverse Events From Elective Surgery Research.
No result was statistically significant on this small set.
[a]The 2-min checkup refers to the daily symptom questionnaire.

Discussion

The SAFER applications were designed to provide timely and accurate information to patients recovering from hysterectomy. This information helped educate patients about normal and abnormal symptoms as well as guide their decision making around accessing health care when experiencing these symptoms. In this report, we focus on the experiences of women who experienced adverse events in our small population, with common ones including infections and wound complications. Although these complications may not have progressed into something more serious, early intervention is critical for successful management. Early, outpatient management of wound complications and infections is less costly to the health care system and patients than unmanaged complications that may progress and require acute care. To this end, the web applications prompted and informed patients on every sign in to remind them about concerning symptoms and, importantly, informing them when and where to go if they had these symptoms.

Mobile app to address barriers to use

A primary goal of this study was to test the feasibility of web-based self-care applications for patients recovering from surgery, complementing standard pre- and post-operative teaching. More than 80 percent of women used the applications regularly, that is, more than half (7) of recommended sign-ins (13). This high usage rate suggests that self-care applications are feasible as a method for helping patients care for themselves at home after surgery. Women who did not use the applications regularly were usually constrained by access to a computer. A recently reported comparison of effectiveness of mobile versus traditional monitoring for weight-loss programs confirms that mobile device versions could be an effective means to increase compliance in some groups.[20] In 2010, 78 percent of Canadian households had a mobile phone, an increasing trend.[21] Data from this small study suggest that a mobile option may help both with physical access and time restrictions by allowing the applications and questionnaires to be completed anywhere. A

mobile phone option may also give greater access to those who live in rural areas without Internet access over land lines but who do have mobile phone coverage. For maximizing compliance with self-management system protocols, easy access from many locations, including bedside, is key. To this end, we have created a mobile-optimized version of the self-care applications which is ready for trial with a larger sample size.

High socioeconomics of small sample size

Most of the participants were relatively young, well educated, and Internet savvy. Selection bias may have occurred in that only women with these characteristics were enrolled or only women with these characteristics had the ability to complete the study because they were required to have home Internet access and familiarity with using the Internet. Due to the social determinants of health, older and low-socioeconomic groups may be at slightly greater risk of adverse outcomes post-surgery. However, a majority of Canadians aged 35–54 (87.8%) and 55–64 (71.1%), as well as low-income (76.2%) Canadians, use the Internet regularly.[22] There is also a demand for relevant, accurate, online health care information among the Canadian female population—74 percent of Canadian female Internet users search for medical or health-related information.[23] Web-based self-care applications can fill this demand by providing patients with reliable, surgery-specific information, reiterating and emphasizing the information provided by their nurses and physicians when patients need it most, namely when they are at home, recovering from surgery, and having problems. The applications also fill the gap after hours and at night when women may feel most vulnerable, have additional questions, and fewer health care professionals are available.

Adverse event rates

Even with the small sample size and the relatively high socioeconomic status of the subjects enrolled in this study, adverse events after abdominal hysterectomy in these subjects were common. Over one-third of participants in this small study experienced an adverse event. Approximately half of these adverse events were infection-related, most commonly urinary tract infections. Large prospective audits and studies with specific measures and questionnaires to capture adverse events unique to hysterectomy also demonstrate adverse event rates of 26.4 and 19.2 percent after abdominal hysterectomy, with most of these adverse events being wound and infection-related.[5,6] With approximately one-fifth to one-third of patients experiencing adverse events after abdominal hysterectomy, the delivery of timely, accurate information about symptoms of possible adverse events and where to go for treatment are needed. Since the application is tailored to the specific day following surgery, the information may be better received and understood because it appears at a relevant time for women seeking the information.

Empowerment of patients

Several aspects of the SAFER website align with previously identified components of patient empowerment.[24] Although the original conception and design of SAFER predates this article (i.e. it was not followed as part of our methods), the daily tip and background information improves communication skills by increasing knowledge of specific questions to pose to health care providers. The 2-min checkup increases knowledge and insight regarding abnormal and normal symptoms, increasing education and health literacy. The daily tip featuring normal and abnormal recovery information specific to the patient's day of recovery gives information about the patient's personal health situation. This advice plus providing daily goals for the patient to strive for, helps

with supporting self-care. The 2-min checkup gives participants specific guidelines of when and where to seek care for flagged symptoms. A final component of patient empowerment identified by the article could be part of future work, namely the encouragement and enabling of having patients communicate with one another. These features strive to help empower patients to care for themselves more effectively.

Additional information and coordination may be required

Although the web applications were helpful for most of the women who experienced an adverse event, one subject who followed the website advice and sought care for her symptoms at local emergency departments still experienced a delay in diagnosis and management of her ureteric injury. This delay was partly from the challenge of diagnosing a small hole in her ureter that presented with non-specific symptoms. There is a need, beyond the currently studied web tool, to co-ordinate surgery-specific self-care information with emergency room physicians and family physicians so that any patients who continue to present to emergency departments or family physicians' offices with ongoing symptoms are referred to appropriate surgical care teams. These teams are experienced in the types of adverse events specific to hysterectomy, which may result in quicker diagnosis, management, and resolution of those adverse events.

Conclusion

Adverse events after abdominal hysterectomy are common. At present, they are not optimally screened and cared for in the Canadian Healthcare system. Self-care applications that are completed regularly while recovering at home after surgery are a feasible method to further educate patients about when and where to go for appropriate evaluation and care beyond the information provided by physicians and nurses at the hospital.

More work still needs to be done with implementation and evaluation—the development of a mobile-optimized app should help with increasing accessibility for recovering patients, which was identified as a barrier to use. Moving from the pilot study stage into a larger trial will allow us to quantify the effect of these applications on adverse event outcomes with statistical support. As we integrate these and other similar applications into health care, we hope to empower patients, decrease adverse event severity, optimize surgical outcomes, and improve the overall health of our population.

Acknowledgements

The authors would like to thank Nicola Sweeney for her assistance with enrolling subjects and collecting data and Melanie Jollymore for her assistance with compiling and revising web-application content for patients.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the Canadian Medical Protective Association Competitive Grant Program, the Canadian Foundation for Innovation, the IWK Research Foundation, and the Nova Scotia Health Research Foundation.

References

1. Spilsbury K, Hammond I, Bulsara M, et al. Morbidity outcomes of 78577 hysterectomies for benign reasons over 23 years. *BJOG* 2008; 115: 1473–1483.
2. Whiteman MK, Hillis SD, Jamieson DJ, et al. Inpatient hysterectomy surveillance in the United States, 2000-2004. *Am J Obstet Gynecol* 2008; 198: 34.e1–34.e7.
3. Gilmour DT and Baskett TF. Disability and litigation from urinary tract injuries at benign gynecologic surgery in Canada. *Obstet Gynecol* 2005; 105: 109–114.
4. Julliard C, Lashoher A, Sewell CA, et al. A national analysis of the relationship between hospital volume, academic center status, and surgical outcomes for abdominal hysterectomy done for leiomyoma. *J Am Coll Surg* 2009; 208: 599–606.
5. Leung PL, Tsang SW and Yuen PM. An audit on hysterectomy for benign diseases in public hospitals in Hong Kong. *Hong Kong Med J* 2007; 13: 187–193.
6. Brummer TH, Jalkanen J, Fraser J, et al. FINHYST, a prospective study of 5279 hysterectomies: complications and their risk factors. *Hum Reprod* 2011; 26: 1741–1751.
7. Perron-Burdick M, Yamamoto M and Zaritsky E. Same-day discharge after laparoscopic hysterectomy. *Obstet Gynecol* 2011; 117: 1136–1141.
8. O'Reilly D. An analysis of perioperative care. *Br J Perioper Nurs* 2001; 11: 402–411.
9. Wade J, Pletsch PK, Morgan SW, et al. Hysterectomy: what do women need and want to know? *J Obstet Gynecol Neonatal Nurs* 2000; 29: 33–42.
10. Lorig KR, Ritter PL, Laurent DD, et al. Internet-based chronic disease self-management: a randomized trial. *Med Care* 2006; 11: 964–971.
11. Fehring RJ, Schneider M and Raviele K. Pilot evaluation of an internet-based natural family planning education and service program. *J Obstet Gynecol Neonatal Nurs* 2011; 40: 281–291.
12. Comer L and Grassley JS. A smoking cessation website for childbearing adolescents. *J Obstet Gynecol Neonatal Nurs* 2010; 39: 695–702.
13. Byczkowski TL, Munafo JK and Britto MT. Family perceptions of the usability and value of chronic disease web-based portals. *Health Inform J* 2014; 20: 151–162.
14. Bunde M, Suls J, Martin R, et al. Online hysterectomy support: characteristics of website experiences. *Cyberpsychol Behav* 2007; 10: 80–85.
15. Bunde M, Suls J, Martin R, et al. Hystersisters online: social support and social comparison among hysterectomy patients on the Internet. *Ann Behav Med* 2006; 31: 271–278.
16. Devine EC. Effects of psychoeducational care for adult surgical patients: a meta-analysis of 191 studies. *Patient Educ Couns* 1992; 19: 129–142.
17. Groves ND, Humphreys HW, Williams AJ, et al. Effect of informational internet web pages on patients' decision-making: randomized controlled trial regarding choice of spinal or general anaesthesia for orthopaedic surgery. *Anaesthesia* 2010; 65: 277–282.
18. Romanov RJ. Building on values: the future of health care in Canada. *Commission on the Future of Health Care in Canada* http://publications.gc.ca/collections/Collection/CP32-85-2002E.pdf (2002, accessed November 2014).
19. Dukeshire S, Gilmour D, MacDonald N, et al. Development and evaluation of a website to improve recovery from hysterectomy. *Comput Inform Nurs* 2012; 30: 164–175.
20. Turner-McGrievy GM, Beets MW, Moore JB, et al. Comparison of traditional versus mobile app self-monitoring of physical activity and dietary intake among overweight adults participating in an mHealth weight loss program. *J Am Med Inform Assoc* 2013; 20: 513–518.
21. Statistics Canada. Residential telephone service survey. *The Daily*, http://www.statcan.gc.ca/daily-quotidien/110405/dq110405a-eng.htm (2010, accessed November 2014).
22. Statistics Canada. Characteristics of individuals using the Internet 2010, http://www.statcan.gc.ca/tables-tableaux/sum-som/l01/cst01/comm35a-eng.htm (2010, accessed November 2014).
23. Statistics Canada. Canadian internet use survey. *The Daily*, http://www.statcan.gc.ca/daily-quotidien/100510/dq100510a-eng.htm (2009, accessed November 2014).
24. Alpay L, Van Der Boog P and Dumaij A. An empowerment-based approach to developing innovative e-health tools for self-management. *Health Inform J* 2011; 17: 247–255.

Appendix 1

Participants were asked to answer "yes" or "no" for the questions below on days 4–10, 13, 16, 20, 23, 27, and 30 after surgery. If they gave a positive reply, they were provided information about the possible significance of the symptom and where to go for assessment. Participants were also given the option of requesting more information about the symptom, without giving a positive reply.

1. Do you have more pain in your abdomen (lower belly) today than you did yesterday?
2. Are the drugs you are taking for pain keeping the pain under control?
3. Do you have pain in your kidney area (i.e. around your lower back ribs)?
4. Are you having any chest pain, upper back pain, or pain in your shoulder tips?
5. Are you short of breath or having any difficulties breathing?
6. Do you have a cough?
7. Do you have any swelling in either one—or both—of your legs?
8. Do you have a fever, chills, or night sweats?
9. Are you lightheaded or dizzy, or have you fainted?
10. Are you having any problems with your incision? For example, is it red, painful, or swollen? Oozing any pus? Opening up?
11. Are you having heavy vaginal bleeding? For example, are you soaking a pad every hour?
12. Is there excessive or foul discharge from your vagina?
13. Have you thrown up in the last 24 h?
14. Are you passing gas rectally?
15. Have you had a bowel movement in the last 24 h?
16. Are you having any diarrhea, or frequent, loose bowel movements?
17. Is there urine or feces (stool) leaking from your vagina?
18. Are you having any problems when you empty your bladder?

Health Informatics Journal

Article

Preserving medical correctness, readability and consistency in de-identified health records

Health Informatics Journal
2017, Vol. 23(4) 291–303

DOI: 10.1177/1460458216647760
journals.sagepub.com/home/jhi
SAGE

Kostas Pantazos*, Soren Lauesen and Soren Lippert
IT-University of Copenhagen, Denmark

Abstract

A health record database contains structured data fields that identify the patient, such as patient ID, patient name, e-mail and phone number. These data are fairly easy to de-identify, that is, replace with other identifiers. However, these data also occur in fields with doctors' free-text notes written in an abbreviated style that cannot be analyzed grammatically. If we replace a word that looks like a name, but isn't, we degrade readability and medical correctness. If we fail to replace it when we should, we degrade confidentiality. We de-identified an existing Danish electronic health record database, ending up with 323,122 patient health records. We had to invent many methods for de-identifying potential identifiers in the free-text notes. The de-identified health records should be used with caution for statistical purposes because we removed health records that were so special that they couldn't be de-identified. Furthermore, we distorted geography by replacing zip codes with random zip codes.

Keywords

anonymity, consistency, correctness, de-identification, electronic health records, readability

Introduction

Electronic health record (EHR) systems store large amounts of data and are essential for all clinical work. According to ANSI,[1] important qualities of an EHR are confidentiality and accessibility only by authorized persons. An EHR system must ensure confidentiality since exposing health records are against law and ethical principles. In order to create data for testing EHR systems, for presenting them to others and for teaching, access is needed to large amounts of EHR data, but it is hard to get the necessary permissions. Access to de-identified (anonymized) health records would in many cases be sufficient. However, the de-identified data should meet certain quality criteria:

*Died October 2015.

Corresponding author:
Soren Lauesen, IT-University of Copenhagen, Rued Langgaards Vej 7, DK-2300 Copenhagen S, Denmark.
Email: slauesen@itu.dk

1. *Medical correctness*: Each health record must show a true medical picture of a real patient.
2. *Anonymity*: It must not be possible to see who the real patient is.
3. *Readability*: The health record must look real. As an example, patient names and addresses that have become *F274, XXXX* or **** don't meet this criterion.
4. *Consistency*: The patient's identifiers must be consistent with the medical picture, for instance, an age that is like the real person's age. If the real patient's name is Peter, but his de-identified name is Jens, then Peter must be replaced by Jens also in the clinician's free-text notes. Furthermore, his wife's health record may refer to him as Peter, and this Peter must also be changed to Jens.

A health record database contains fields with structured data that can identify the patient, such as patient ID, patient name and phone number. These data are fairly easy to replace with other identifiers, in that way ensuring anonymity. The database also contains fields with medical data such as diagnosis codes, blood pressure and other measured values. They have to be preserved to ensure medical correctness. The problem is the unstructured data fields with doctors' free-text notes. They contain important medical information that has to be preserved, but may also contain phone numbers, patient names, name of the spouse and other identifying items. Furthermore, clinicians write notes in an abbreviated—often personal—style that cannot be analyzed grammatically.

Considerable work[2–12] has been done in developing de-identification algorithms using various techniques such as natural language processing (NLP), named entity recognition and machine learning. These approaches de-identify database records (e.g. pathology documents) that do not relate to other records. We will refer to this as a record-oriented de-identification approach. Recent work[13–15] has focused on utilizing a full database rather than records. The approach presented in this article was briefly presented by Pantazos et al.[16]

Previous research has not looked at quality attributes for database-oriented de-identifications. In this article, we focus on the four quality attributes above: medical correctness, anonymity, readability and consistency.

Background

In 2010, we started work on an EHR system with a high degree of data visualization. We cooperated with a Danish software house that had delivered EHR systems to many clinics and small hospitals in Denmark. In order to get test data, we made a copy of the full database and de-identified it. The database consisted of 437,164 patient health records. The work took place on their premises since no real health records could go outside the company.

The idea was to make a mapping table that translated all patient identifiers into patient identifiers for other patients. In this way, patient B got patient C's first name, patient D's last name, patient E's street name and so on. Patient B would get a randomized civil registration number (CPR) that preserved his year of birth and gender. In this way, we would ensure consistency across patients. Somebody looking at a full de-identified patient record would know that this was a real patient, but he or she was not called C, nor D and didn't have address E. We had outlined the conversion program and expected the whole thing to take a couple of days, but—alas—unexpected problems turned up. We spent 3 months.

We had to invent many methods for locating and anonymizing potential identifiers in the free-text notes. To our surprise, 3–4 percent of the words in free-text fields were potential patient identifiers. Consistency across patients turned up to be more important than we had expected: around 90 percent of the patients had one or more relatives in the database.

We detected that many health records had been created for testing or were left uncompleted due to system errors ("corrupt" data). In 69,914 cases, we had to delete such patient records. In 43,119 cases, data couldn't be safely de-identified without manual intervention, so we made the program delete these patient records. As an example, we deleted patients that had a rare Danish name that was also the name of a disease, for instance, Aaron, which is also a medical term. The program couldn't tell whether Aaron in a free text was a medical term that shouldn't be changed or the name of a related patient that should. We ended up with 323,122 patient health records.

We manually compared 369 random, anonymized patient records against the original records, checking for medical correctness, readability and anonymity. These quality factors were preserved on an acceptable level for our purpose. Consistency was ensured by the algorithm, so we checked it in a few places only.

Related work

A good de-identification system must replace all data that are personal identifiers in structured data, as well as in free text.[5] One of the first de-identification systems for patient records was Scrub.[2] It was evaluated against 275 English patient records and 3198 letters to physicians from the pediatric department. External sources, predefined templates and rules (e.g. the format of a phone number and address) were included in the algorithm. This algorithm had a 99–100 percent success rate for de-identifying personal identifiers. Another system was developed by Ruch et al.[3] It resembles Scrub, but added NLP. NLP tools use a medical semantic dictionary with word-sense and morph-syntactic labeling. This system located 98–99 percent of all personal identifiers. To anonymize the data, the authors replaced all identifiers with XXX's, which had a negative impact on readability.

Several systems[4–6,17] were developed in the last decade to de-identify pathology reports. Thomas et al.[4] developed an algorithm that scored 98.7 percent successful name replacements using English syntactic rules, prefixes, suffixes and names composed of first and last names. Gupta et al.[6] conducted an iterative evaluation of their system. At the end of the third iteration, the authors claimed that their method generated anonymized and readable reports. An algorithm designed by Berman[5] replaced words with codes from the Unified Medical Language System (UMLS) and asterisks. It produced hardly-readable documents. Beckwith et al.[17] evaluated an open source system which replaced identifiers with X's in pathology reports.

Seven de-identification systems were evaluated in the "Challenge in NLP for clinical data" workshop, using medical discharge letters as input.[18] In this workshop, the systems were evaluated using three performance measures: precision, recall and f-measure. The highest f-measure was 99.7534 achieved by a novel approach based on Named Entity Recognition combined with iterative machine learning.[8] This application finds personal identifiers in the structured data and uses them to locate identifiers in free-text data.

Hanauer et al.[12] introduced the iterative tag-a-little, learn-a-little approach for a particular document type. The authors used the MITRE Identification Scrubber Toolkit[19] to integrate their approach. They obtained an f-measure of 95.

Susilo and Win[20] present a new approach for patient confidentiality that utilizes searching through encrypted data. Huang et al.[21] focused on portable EHRs for privacy preservation. The authors stress the feasibility of the approach, which can meet patient confidentiality requirements.

Even though most of the research has been in an English context, there are some studies on de-identifying in other contexts. Tveit et al.[22] present their approach to de-identify Norwegian general practitioner medical records. Their approach consists of six steps: create dictionaries, find exact match and tag, identify approximate match and tag, replace tags, tackle untagged words and

generate the de-identified output. However, this approach was not evaluated empirically. A Swedish de-identification system was developed by Kokkinakis and Thurin,[7] using named entity recognition. This approach de-identified 200 Swedish discharge letters with a precision of 96.97 percent, recall of 89.35 percent and f-measure of 93. Velupillai et al.[10] adjusted an English de-identification system for Swedish medical records. This transformation did not produce the expected results (f-measure in total = 65, f-measure for names = 80). Consequently, the authors reported that building a Swedish system from scratch was more efficient. This phenomenon was also observed and confirmed by Grouin et al.[9] who adjusted a de-identification system from English to French and obtained poor results.

Meystre et al.[23] reviewed recent de-identification algorithms and found that the majority of the algorithms focus on de-identifying structured data and not free text. However, in accordance with Dalianis and Velupillai[24] and Hanauer et al.,[12] there is immense valuable information in the free text. We found the same in our data.

Quality factors

An EHR contains database fields with structured data that can identify the patient, for instance, CPR, patient name and phone number. Other structured data fields contain medical data such as diagnosis codes and blood pressure. The EHR also contains free-text fields, for instance, doctor's notes and discharge letters. It may also contain pictures of body parts, X-ray and so on, usually with a patient ID embedded in the picture. We have not dealt with pictures in this project.

Some data are *quasi-identifiers* because they can narrow down the set of patients that might have this health record. Examples are street name, zip code, birth date, hospital or clinician who treated the patient. Two or more quasi-identifiers in combination may identify the patient.[25]

Anonymity

In order to ensure anonymity, all patient identifiers and quasi-identifiers must be de-identified, that is, replaced with something else. It is fairly easy to do this for structured data, but very hard for free-text data. Often the computer has no way to tell whether a free-text word is an everyday word, a medical term or part of a patient name. As an example, Aaron's sign is a medical term, but it might also be the name of a person.

Readability

In order to ensure readability, we have to replace the patient name with a new name that looks real. Inside this patient's record, we have to be consistent so that we replace with the same name for all occurrences.

Consistency

In the database we worked with, 90 percent of the patients had one or more relatives in the database. Most likely, the patient's name and/or CPR will occur in one of these related health records in free-text fields. To ensure consistency, we have to replace also these identifiers with the same new identifier.

There are other aspects of consistency, for instance, that the distribution of names should remain much the same. If rare names suddenly turn up for a large number of patients, the health record database will not look real.

Medical correctness

If we replace a medical term that looks like a person name, with the new person name, the health record will look odd. We have lost medical correctness and readability. In many cases, a clinician can guess what the medical term was and in that way get to know the original name of all patients that have this new name.

Another aspect of medical correctness is age. If birth dates are transformed in a way that makes the patient have a very different age, it will not match the patient's diagnosis pattern.

Solution

We will first give an overview of the solution and then explain the details and where the data came from.

Permutation tables

For some identifiers, we made a *permutation table* that mapped existing identifiers to new ones. We picked the new identifier at random from the same table, avoiding reuse of identifiers. Any occurrence of an identifier from this table would be translated into the new one.

As an example, we created a permutation table of all last names. The last name Jensen would be translated into Petersen wherever it occurred. Petersen was another last name in the table, with a similar frequency. This ensured readability and consistency across all patient records.

We made permutation tables for these identifiers and quasi-identifiers: first male names, first female names, last names, street names, zip codes, hospital and clinic names.

Distorted identifier table

For the CPR, we made a mapping table from existing CPR to a distorted CPR in this way: The Danish CPR format is: DDMMYY-CSSG where DDMMYY is the birth date. The day (DD) and month (MM) were changed to a random valid day and month. The year (YY) was kept. C indicates birth century (1900 or 2000). This was not changed. SS (serial number) was changed, while G indicates the gender and wasn't changed.

This ensured readability (clinical users see lots of CPR numbers and can easily spot wrong ones) and medical correctness (because age and gender were kept).

Randomized identifier

For other identifiers, we randomized the identifier without caring about readability or consistency. This applied to phone numbers, e-mail addresses and URLs.

Ambiguous words

Ambiguous words could be part of a person's name or something else, for instance, a medical term or a common word. Through many sources, we created a list of ambiguous words. When the de-identification program meets an ambiguous word B in a free-text field, it has three choices:

1. Replace the word B with its corresponding new name, C. If the word B actually is part of a person's name, everything is fine. But if B actually is a medical term or a common word,

the clinician can see from the context that C probably means B. If he knows the replacement rules, he now knows that everybody in the database with the name C is actually B. We lose not only medical correctness and readability but also some anonymity. For this reason, we never replace ambiguous words.

2. Keep the old word B. This ensures medical correctness, readability and consistency. If the word actually is a person's name, the clinician can see it from the context. As an example, assume that the program finds Aaron in a free-text note. Since it is a medical term, it keeps it. However, a clinician can see that Aaron in this context is the name of a person. If he knows the rule of replacement, he now knows that the person referred to is really called Aaron, although this is not his name in the de-identified database. The clinician gets no clue to where Aaron's health record is. If there are only a few Aarons in real life, he might guess whom it is. If there are many Aarons, he cannot know. We decided that 200 occurrences was a safe limit. If the ambiguous name occurs more than 200 times, we keep it in the database.
3. Delete all patient records with name B. We do this when the ambiguous name occurs less than 200 times. This ensures all four quality factors, but we lose data. If a free text for another patient refers to patient B, the reference will now be to a deleted patient. We have lost a bit of consistency, but such data could exist anyway in the database.

The database and the mapping tables

The EHR we de-identified is built on Microsoft Axapta, which is an ERP system that can be extended in many ways. It contained data from 79 clinics and hospitals (including a few in Greenland and the Faroe Islands) and contained 437,164 patient records in total. The entire database was 12 GB. There were 65 health-related tables:

1. 43 tables had no fields that could expose the patient identity. They included reference tables of drug codes, treatment codes and diagnosis codes.
2. 9 tables had fields that only contained personal identifiers in structured form, for instance, the patient table that contained patient ID, first name, last name, address, zip code, five phone numbers, birth date and date of death. Another example is a table of family relations, that is, relations between two patients. Clinicians, hospitals and clinics had their own tables with name, address and so on.
3. 13 tables had fields with free text. The largest one was Medical Record Lines, which occupied 7 of the 12 GB in the database.

Mapping tables

To be able to replace existing identifiers with new identifiers, we created the following mapping tables.

CPR. We collected the CPR numbers from the patient table and gave each number a partially random new number according to the rules above. If the new number was already used as a new number, we randomized it once more.

Last names. We used three sources to collect last names: the database's patient table, Danmarks Statistik's website[26] and a study of Danish names at University of Copenhagen, 2005.[27] We merged

these sources and obtained 56,339 last names. We counted how often each last name occurred in the patient table. Many names didn't occur at all in the patient table, but might occur in free-text notes. It was important to catch them too and de-identify them.

For each name, we assigned a new name from the table with a frequency similar to the old name. We used this approach: we divided the names into groups according to frequency. The group of most frequent last names consisted of 20 names with frequencies from 14,712 to 5319. We rotated these names a random number of steps to obtain the new names (a cyclical permutation). We used the same approach for groups of 30 names with decreasing frequencies. Rare names (frequency < 200) were randomly replaced with another name in the frequent part of the list. This also took care of the names that didn't occur at all in the patient table.

Male first names. We used the same approach to collect and de-identify male first names. For names occurring in the patient table, we got the gender from the CPR number. Our external sources had separate lists for male and female first names. In total, we got 11,415 male first names.

Female first names. We treated them in the same way as male names. In total, we got 13,044 female first names.

Street names. We collected street names from the patient table's address field. The address field included also floor numbers and entrance letters. In Denmark, the street name is first, so we simply extracted the first real name from the address field. We also included street names from the CPR website. In total, we got 25,429 street names. We assigned a random street name as the new name without caring about frequencies or consistency with zip codes.

Zip codes. We collected zip codes and related city names from Post Danmark[28] and assigned a random zip code and city name as the new name. In total, we got 1396 zip codes.

Hospital names and clinic names. We collected hospital names from Region Hovedstaden, Region Sjælland, Region Syddanmark, Region Midtjylland, Region Nordjylland, Queen Ingrid's Hospital in Greenland, Faroe Islands website and our own EHR Database. We used Sygehusvalg,[29] Brancheforeningen for Privathospitaler og Klinikker (the trade association) and our own EHR database to extract names of clinics. In total, we got a list of 219 clinic names and 93 hospital names. We did not randomly assign new names to the clinics and hospitals. This would reduce medical correctness because clinicians know which clinics do what. On the other hand, being treated in a specific clinic is a quasi-identifier. We manually selected 41 hospital names and 92 clinic names and used them as new names. In many cases, the new name was simply "Hospital" or "Clinic." This was a reduction in readability and to some extent in medical correctness.

Ambiguous names

Ambiguous names in our context are first or last names of persons that happen to mean something else too. We need a table of them to decide how to treat such a name when it occurs in free text. As explained above, we have to delete patients with rare names if they appear in free text. If they are frequent names, we leave them as they are.

In healthcare, it is common that diseases, signs, symptoms and so forth are named after a person, most likely the one who discovered it. These names are called medical eponyms and may cause ambiguity. For example, according to Statistics Denmark in 2010,[26] there were 88 males using the

name Aaron. At the same time, Aaron is part of a medical eponym (Aaron sign). The algorithm knows too little about the context to decide whether to de-identify this name or not.

A similar ambiguity exists also with common words in a language. Each language contains several words whose meaning depends on the context. For example, in Danish, the word "hans" can be a pronoun or a male name. Another ambiguous case is abbreviations used by clinicians. For instance, instead of writing "kirurgisk" (in English: "surgical") they use the abbreviation "kir," which can be a last name as well. As another example, it is common that a city, hospital, clinic or street name is used as a first or last name. For instance, Aalborg is a city in Denmark, but it is a last name as well.

We derived the table of ambiguous names from several sources. We checked our lists of first and last names against the Danish Dictionary from Microsoft Office Word 2010. This created a list of 3557 potentially ambiguous names. That a name exists in the dictionary doesn't mean that it also has another meaning that can occur in health records. So, the medical specialist in our team (Lippert) scrutinized the list and came up with 1952 ambiguous names.

To the best of our knowledge, there is no official source that contains medical eponyms. So, we used the website "Who named it"[30] and extracted 3246 medical eponymous names from it. They too entered the list of ambiguous names.

Applying the mappings

The mappings must be applied to the structured fields as well as to the free-text fields. We applied the mappings to the structured fields according to Table 1. Notice the last rule: remove all patients above 90 years. It came from the US Health Insurance Portability and Accountability Act guidelines, HIPAA.[31] There are so few patients above 90 that their age exposes them.

It was harder to apply the mappings in free-text fields because we don't know whether an identifier is a first name, a street name and so on. The program analyzed the free-text token by token and applied these rules.

Name tokens

1. If the name is in one of the person name tables and also in the table of ambiguous words, do nothing or delete the related patient records depending on the name frequency.
2. If the name is in the last name table, replace it with the new name in the table.
3. If the name is in the first male name table, replace it with the new name.
4. If the name is in the first female name table, replace it with the new name.
5. If the name is in the table of ambiguous words, leave it as it is.
6. If the name is in the table of street names, replace it with the new name.
7. If the name is in the table of zip codes and city names, replace it with the new city name.
8. If the name is in the table of hospitals and clinics, replace it with the new name.
9. Otherwise, leave it as it is.

Number tokens

10. If the number is in the table of CPR numbers, replace it with the new CPR number.
11. If the number looks like a CPR number (10 digits starting with a date), randomize it as other CPR numbers.
12. If the number has eight digits and is next to a word like tlf, tel and fax, randomize the number.

Table 1. Replacement rules for structured data fields.

Identifying fields	
Civil registration number (CPR)	Replace it with the new CPR in the CPR mapping table
First name	Select the first male or first female mapping table according to the gender code in CPR. Replace first name with the new name in the mapping table
Last name	Replace it with a new name according to the mapping table
Address	An address contains a street name, a house number and sometimes a floor number and entrance position (e.g. Byvej 21, 2tv). Replace the street name according to the street mapping table. Replace numbers randomly with a number that has the same number of digits
Phone numbers (up to five per patient)	Alter each phone number to a random number with the same number of digits
E-mail	Alter the address with random characters before the letter @ and change the domain name to *email.dk*
Quasi-identifiers	
Zip code	Replace it according to the zip mapping table
City	Replace it with the city name in the zip mapping table
Country	Change it to *Denmark*
Date of birth	Set it from the new CPR
Date of death	Randomly change the day and month
Hospital name	Replace it with a new name according to the mapping table
Clinic name	Replace it with a new name according to the mapping table
Clinician first name	Replace it with a new name according to the mapping table for first names
Clinician last name	Replace it with a new name according to the mapping table for last names
Clinician alias	Replace it with the new first name of the clinician
Age	Remove all patients older than 90 years due to high anonymity risks

13. If the number is in the table of zip codes and next to a city name, replace it with the new zip code.
14. Otherwise, leave it as it is. (It may be a measured value, a lab-test number (eight digits), a house number and so on.)

These rules give priority to anonymity rather than medical correctness. As an example, a lab-test number or a date-time that looks like a CPR number will be de-identified and thus reduce the medical correctness.

Evaluation of the quality factors

Anonymity, readability and medical correctness

In order to evaluate the actual anonymity, readability and medical correctness, we need to know how many words were replaced incorrectly.

We selected a random sample of 369 full patient records. A clinician manually compared all the free-text fields in the old and the new version, in total 73,150 words. The result is shown in Table 2.

Table 2. Correct and incorrect replacements.

Number of words	Should be de-identified	Should not	Total
Was de-identified	1313	109	1422
Was not	7	71,721	71,728
Total	1320	71,830	73,150

Seven words should have been de-identified but wasn't. Only one of them was a person name. It was ambiguous and frequent (frequency >200) and consequently preserved according to our rules. Since it was frequent, we consider it a quasi-identifier. The other words were quasi-identifiers such as department names and misspelled street names that were not in our translation tables. In total, out of 73,150 words, we had seven anonymity leaks on quasi-identifiers and none on full identifiers.

A total of 109 words were replaced, but shouldn't. They were ambiguous, but not in our table of ambiguous words. One example was the word "Uno," which was the name of a drug, but also a male first name. These cases decreased medical correctness and readability. It also revealed a general weakness: also drug names should be considered a source of ambiguity.

Measured in the traditional way with *recall* and *precision*, the algorithm scored 99.5 percent for recall (the seven anonymity leaks) and 92.3 percent for precision (the 109 leaks in medical correctness). The f-measure was 95.7 percent. Our database-oriented approach compares favorably with previous work on record-based de-identification approaches.[4–6,17]

It would be interesting to compare how other de-identification approaches would handle our data. However, this is impossible because the approaches are very dependent on the language. Furthermore, we are not allowed to move our original data out of the company where it is hosted. We have not found publications about de-identification that discuss ambiguity. Most likely, they don't pay attention to it. It will probably cause some leaks of confidentiality that isn't detected.

Consistency

The database can record family relations and other relations between patients. Around 90 percent of the patients have one or more recorded relatives. When a person name is de-identified in the structured patient table, it is important that the same name is de-identified in the same way in the rest of the patient's records and in records of related patients, also for free-text fields. This is solely a matter of correct programming. We checked it for a couple of patients in Table 3. Since the translation tables are used for all patients, consistency is also preserved for relatives who are not recorded as relatives.

Results

Table 3 shows a (non-random) sample of patients with two or more relatives. It gives an impression of the variety and complexity of patient records. Several patients have eight relatives in the database, many have more than 100 measurements with notes (Clinical Data), many have more than 10 diagnoses and several hundred prescriptions.

Table 4 shows the results for the Medical Record Line and Clinical Data tables, which contain most of the free text in the database. In total, 3–4 percent of the words are personal identifiers.

This study is the first de-identification algorithm that focuses on anonymity, medical correctness, readability and consistency. Other approaches are limited to a few types of documents, while our approach deals with full EHR records from 79 hospitals and clinics. An important part of our approach was to collect ambiguous names from many sources.

Table 3. Sample of 20 patients showing the variety of patient records.

CPR	Relatives	Clinical data	Medical records	Diagnoses	Prescriptions	Total
2905931069	6	54	4	4	6	68
2904220702	2	335	9	3	678	1025
2812620120	4	37	2	42	177	258
2812351528	2	36	1	54	517	608
2811831753	2	30	1	1	18	50
2810291211	2	22	2	13	68	105
2809711115	4	15	4	9	15	43
2809550048	6	151	6	2	32	191
2808972414	8	50	6	9	7	72
2806492477	4	603	10	22	412	1047
2805832168	4	42	1	11	62	116
2805620030	4	176	5	1	29	211
2803961559	2	6	3	2	5	16
2801981465	8	76	4	1	1	82
2801460257	6	29	3	1	22	55
2712742278	6	186	8	7	64	265
2711743812	4	77	9	14	51	151
2711440133	2	98	4	11	100	213
2710592476	8	238	3	10	38	289
2709530059	4	22	1	1	9	33

CPR: civil registration number.

Table 4. Number of identifiers in free text.

	Medical record line	Clinical data
E-mails	18,858	727
Phone numbers	43,051	62,461
Clinics	114,318	17,213
CPRs	455,946	121,036
Zip codes	599,566	668
Hospitals	787,055	117,369
Cities	994,125	7557
Last names	2,675,386	254,915
Street names	3,156,356	125,470
First names	4,331,593	330,679
Total identifiers	(4%) 13,176,254	(3%) 1,038,095
Non-identifiers	322,734,954	32,052,044

CPR: civil registration number.

We started out with 437,164 patient health records. We deleted 69,914 patient records because data were corrupted (old test data and records left after system failures). We deleted 43,119 patient records because of rare ambiguous names or because the patient was older than 90. We ended up with 323,122 patient health records.

The distinction between frequent and rare names (fewer than 200 occurrences) is somewhat arbitrary. The limit of 200 caused us to delete "only" 43,119 patient records because they had rare ambiguous names. If all names were considered rare, we would have lost another 55,000 patient records.

We made a manual review of 369 patient records with 71,721 free-text words. It revealed seven words where a quasi-identifier hadn't been de-identified. It revealed 109 words where it was de-identified, but shouldn't because the word wasn't in our list of ambiguous names. This reduced medical correctness and readability.

Limitations and errors

An EHR database contains also binary files (e.g. X-rays), scanned documents and Word documents. They are not part of the database, but fields in the database contain the file names. Our approach is limited to structured data and free-text fields, and it doesn't try to de-identify pictures and other files. The picture will usually contain patient identifiers such as CPR and name. De-identifying these would be a project of its own.

We have not tried to deal with spelling errors. It might have reduced the seven un-identified words above to around three. We could deal with spelling errors by looking at close matches of words instead of precise matches, but we don't know how much it would have increased the number of false de-identifications (the 109 words above).

We forgot to put also pharmaceutical names in the list of ambiguous words. This could have removed some of the 109 false de-identifications above.

We missed several clinical abbreviations as potential ambiguous names. A language analysis of the free-text notes might have revealed them.

The de-identified data should be used with caution for statistical purposes because of the way we had to remove health records that couldn't be de-identified and also because we deleted patients older than 90 and distorted geography by replacing zip codes with random zip codes.

For statistical purposes, the de-identification should have been different. We shouldn't care about readability or consistency, but simply replace all potential identifiers in free text with asterisks or the like. We should only delete corrupted patient records. The mapping tables would still be needed, but only to detect what might be an identifier. We wouldn't need to care about ambiguous words. The result would probably be similar to many other de-identification approaches.

Funding

The author(s) received no financial support for the research, authorship and/or publication of this article.

References

1. ISO/TR 20514:2005. Health informatics—electronic health record definition, scope and context.
2. Sweeney L. Replacing personally-identifying information in medical records, the scrub system. In: *Proceedings of the AMIA annual fall symposium*, Washington, DC, 1996, p. 333. Bethesda, MD: American Medical Informatics Association. Available at: http://dataprivacylab.org/projects/scrub/paper1.pdf
3. Ruch P, Baud RH, Rassinoux AM, et al. Medical document anonymization with a semantic lexicon. In: *Proceedings of the AMIA symposium*, Los Angeles, CA, 4–8 November 2000, p. 729. Bethesda, MD: American Medical Informatics Association.
4. Thomas SM, Mamlin B, Schadow G, et al. A successful technique for removing names in pathology reports using an augmented search and replace method. In: Proceedings of the AMIA symposium, San Antonio, TX, 9–13 November 2002, p. 777. Bethesda, MD: American Medical Informatics Association.
5. Berman JJ. Concept-match medical data scrubbing: how pathology text can be used in research. *Arch Pathol Lab Med* 2003; 127(6): 680–686.
6. Gupta D, Saul M and Gilbertson J. Evaluation of a de-identification (de-id) software engine to share pathology reports and clinical documents for research. *Am J Clin Pathol* 2004; 121(2): 176–186.

7. Kokkinakis D and Thurin A. Identification of entity references in hospital discharge letters. In: Proceedings of the 16th Nordic conference of computational linguistics, Tartu, 25–26 May 2007, pp. 329–332.
8. Szarvas G, Farkas R and Busa-Fekete R. State-of-the-art anonymization of medical records using an iterative machine learning framework. *J Am Med Inform Assoc* 2007; 14(5): 574–580.
9. Grouin C, Rosier A, Dameron O, et al. Testing tactics to localize de-identification. In: K-P, et al. (eds) *Medical Informatics in a United and Healthy Europe* vol.150. Amsterdam, The Netherlands: IOS Press, 2009, pp. 735–739.
10. Velupillai S, Dalianis H, Hassel M, et al. Developing a standard for de-identifying electronic patient records written in Swedish: precision, recall and f-measure in a manual and computerized annotation trial. *Int J Med Inform* 2009; 78(12): e19–e26.
11. Dalianis H, Nilsson G and Velupillai S. Is de-identification of electronic health records possible? Or can we use health record corpora for research? In: AAAI fall symposium series, Arlington, VA, 5–7 November 2009.
12. Hanauer D, Aberdeen J, Bayer S, et al. Bootstrapping a de-identification system for narrative patient records: cost-performance trade-offs. *Int J Med Inform* 2013; 82(9): 821–831.
13. Emam KE, Paton D, Dankar FK, et al. De-identifying a public use microdata file from the Canadian national discharge abstract database. *BMC Med Inform Decis Mak* 2011; 11: 53.
14. Viangteeravat T, Huang EY and Wade G. Giving raw data a chance to talk: a demonstration of de-identified pediatric research database (PRD) and exploratory analysis techniques for possible research cohort discovery and identifiable high risk factors for re-admission. *BMC Bioinformatics* 2013; 14: A5.
15. Gordon JS. Altering the function of the electronic medical record: creating a de-identified database for clinical researchers and educators. *Nurs Inform* 2012; 2012: 132.
16. Pantazos K, Lauesen S and Lippert S. De-identifying an EHR database—anonymity, correctness and readability of the medical record. *Stud Health Technol Inform* 2011; 169: 862–866.
17. Beckwith BA, Mahaadevan R, Balis UJ, et al. Development and evaluation of an open source software tool for deidentification of pathology reports. *BMC Med Inform Decis Mak* 2006; 6(1): 12.
18. Uzuner O, Luo Y and Szolovits P. Evaluating the state-of-the-art in automatic de-identification. *J Am Med Inform Assoc* 2007; 14(5): 550–563.
19. Aberdeen J, Bayer S, Yeniterzi R, et al. The MITRE identification scrubber toolkit: design, training, and assessment. *International J Med Inform* 2010; 79(12): 849–859.
20. Susilo W and Win K. Security and access of health research data. *J Med Syst* 2007; 31(2): 103–107, http://dx.doi.org/10.1007/s10916-006-9035-y (accessed August 2010).
21. Huang LC, Chu HC, Lien CY, et al. Embedding a hiding function in a portable electronic health record for privacy preservation. *J Med Syst* 2010; 34(3): 313–320, http://dx.doi.org/10.1007/s10916-008-9243-8 (accessed August 2010).
22. Tveit A, Edsberg O, Raast TB, et al. Anonymisation of general practitioner's patient records. In: Proceedings of the HelsIT'04 conference, Trondheim, September 2004.
23. Meystre SM, Friedlin FJ, South BR, et al. Automatic de-identification of textual documents in the electronic health record: a review of recent research. *BMC Medical Res Methodol* 2010; 10(1): 70.
24. Dalianis H and Velupillai S. De-identifying Swedish clinical text-refinement of a gold standard and experiments with conditional random fields. *J Biomedical Semantics* 2010; 1: 6.
25. El Emam K, Jabbouri S, Sams S, et al. Evaluating common de-identification heuristics for personal health information. *J Med Internet Res* 2006; 8(4): e28
26. Danmarks Statistik, http://www.dst.dk/HomeUK/Statistics/Names.aspx (accessed August 2010).
27. Copenhagen University, http://danskernesnavne.navneforskning.ku.dk/TopNavne.asp (accessed August 2010).
28. Post Denmark, http://www.postdanmark.dk/da/Privat/Kundeservice/postnummerkort/Sider/postnummerkort.aspx (accessed May 2016).
29. Sygehusvalg, http://www.sygehusvalg.dk/ (accessed August 2010).
30. Who Named It, http://www.whonamedit.com/azeponyms.cfm/ (accessed August 2010).
31. HIPAA. HIPAA privacy rules and public health, http://www.cdc.gov/mmwr/preview/mmwrhtml/m2e411a1.htm (accessed August 2010).

Health Informatics Journal

Article

The psychosocial effect of web-based information in fast-track surgery

Health Informatics Journal
2017, Vol. 23(4) 304–318

DOI: 10.1177/1460458216648477
journals.sagepub.com/home/jhi
SAGE

Lene Bastrup Jørgensen
Aarhus University, Denmark; Regional Hospital Silkeborg, Center for Planned Surgery, Denmark

Lone Ramer Mikkelsen
Regional Hospital Silkeborg, Center for Planned Surgery, Denmark

Bodil Bjørnshave Noe
Regional Hospital Herning, Center for Research and Education, Denmark

Martin Vesterby
Aarhus University, Denmark

Maria Uhd
Regional Hospital Silkeborg, Center for Planned Surgery, Denmark

Bengt Fridlund
Jönköping University, Sweden

Abstract

The psychosocial effects of web-based information have yet to be tested for patients joining a fast-track total hip arthroplasty programme. This study compared and evaluated the psychosocial impact of standard total hip arthroplasty programme, with and without supplementation with a web-based information platform (E-total hip arthroplasty programme). Totally, 299 patients were enrolled in an un-controlled, before-and-after study, 117 in the S-total hip arthroplasty programme group and 182 in the E-total hip arthroplasty programme group. Psychosocial outcomes before and during admission and then 3 months post-surgery were evaluated, with analyses conducted between and within groups. All outcomes improved significantly from pre-admission to 3 months post-surgery, with no between-group differences. In all, 112 of the 182 E-total hip arthroplasty programme patients accessed the learning platform. A subgroup analysis showed no

Corresponding author:
Lene Bastrup Jørgensen, Interdisciplinary Research Unit, Center for Planned Surgery, Regional Hospital Silkeborg, Falkevej 1-3, 8600 Silkeborg, Denmark.
Email: lenejoer@rm.dk

significant differences between users and non-users, either at baseline or in terms of outcome. This study found no positive psychosocial effect between groups, but a significant effect within groups.

Keywords
fast-track, intervention, orthopaedic patients, psychosocial state, web-based animated information

Introduction

Web-based health informatics using animated information has the potential to reduce the complexity of preoperative information, thereby improving the mental health of patients undergoing surgery. However, this expectation has yet to be tested for patients accessing a fast-track total hip arthroplasty (THA) programme, which presupposes comprehensive educational and mental resources.[1,2] The most common indication for THA surgery is arthrosis, with complaints of hip pain and reduced functional ability.[3] Demand for THA procedures is expected to increase in the future as a function of greater physical activity, together with an expectation among elderly patients of an improved quality of life.[4]

In 2001, Wilmore and Kehlet[5] were the first surgeons to describe the concept of fast-track treatment. Later, they described fast-track regimes as a treatment which 'focuses on enhancing recovery and reducing morbidity by implementing evidence in the fields of anaesthesia, analgesia, reduction of surgical stress, fluid management, minimal invasive surgery, nutrition, and ambulation'.[6,7] Thus, fast-track treatment focuses on optimising the preoperative education of patients, attenuating their surgical stress response, optimising pain relief, enforcing mobilisation, nutritional support and the provision of up-to-date postoperative nursing care and rehabilitation.[8,9] In this way, a fast-track programme aims to optimise recovery, reduce hospital stays and diminish postoperative morbidity.[7,10] Patients are encouraged to take an active part in their treatment and rehabilitation and to use the fast-track total hip arthroplasty programme (THAP) as an accelerated intervention that commences during the preoperative period (with an information meeting scheduled prior to admission) and continues through hospitalisation until the day of discharge.[11]

Emotional health among patients going through orthopaedic surgery has an influence on physical recovery post-surgery.[11] Preoperative anxiety is a well-known and recognised mental state which influences cognitive ability.[12,13] Reducing anxiety may therefore help to improve learning ability,[14] and improving informational level and mode may reduce anxiety.[14,15] Thus, learning capacity and mental health are mutually reinforcing, a consideration which should be 'factored-in' to current practice in health education.

In line with this, a review on computerised cognitive behavioural therapy (CCBT) has documented success by providing patients with non-catastrophic images of anxiety-triggering processes.[16] Using this technique, animated health information seems to confer some benefit in terms of preoperative anxiety.[17] The visual approach optimises the acquisition of knowledge by reducing data complexity,[18–21] as well as minimising cultural and personal differences. Animation using advertising techniques such as light, colour, size and music also confers beneficial effects in terms of attentiveness to the displayed material and knowledge retention.[18,20,21]

Thus, web-based and animated information may help to bridge unmet needs in health literacy,[22,23] given that both media types improve learning irrespective of health literacy.[20,24] Furthermore, strengthening health literacy correlates positively with improved learning ability and a high level of self-efficacy in accomplishing postoperative tasks.[25,26]

The beneficial impact of pre-surgery animated health information and Internet-based education on patients' health, especially mental health as well as learning ability, has yet to be evaluated for patients

undergoing a resource-intensive fast-track THAP. Furthermore, the identification of clinically relevant predictors of patient anxiety in relation to THAP can facilitate the provision of individualised support to the most susceptible patients. Very few studies have identified such characteristics with a view to their possible influence on fast-track pre- and postoperative processes.[22]

Objectives

Our objectives were twofold. First, to compare and evaluate the impact of a standard THAP (S-THAP) with an S-THAP extended with a web-based information platform (WIP) (E-THAP) with regard to chance in anxiety (primary outcome), depression, health condition and self-efficacy from pre-surgery to 3 months post-surgery.

Second, to determine whether anxiety pre-surgery is a predictor to both anxiety during hospital stay and anxiety, depression, health condition and self-efficacy 3 months post-surgery.

Methods

Design and setting

Our study incorporated an un-controlled before-and-after design and was hosted from September 2012 to October 2014 at an elective surgical centre in a Danish hospital, which performs approximately 700 THA procedures annually. A fast-track THAP, in accordance with the fast-track criteria described by Kehlet and colleagues,[5,7] is practised at this centre.

THAPs

For S-THAP, the expectation is that patients are discharged the day after surgery, which leaves relatively little time in which to prepare them for surgery, discharge and post-surgical rehabilitation. In general, the time frame from setting the date of surgery in the outpatient clinic, to the day of surgery, was approximately 2–3 weeks. Patients were informed about their surgery at the outpatient clinic by a surgeon, and then by a specialist nurse. The nurse also provided a 36-page pamphlet to patients who were advised to read it thoroughly before attending the information group meeting, 1–2 weeks prior to surgery. Patients as well as relatives were invited to this 2-h meeting, which included oral and written information, together with a power-point presentation of the THA procedure and postoperative recovery. A clear division of pre- and post-surgery tasks between patients, their relatives and healthcare professionals was emphasised as a prerequisite for achieving a successful surgical outcome.[27] Throughout the programme, starting from the day of their outpatient clinic appointment, and ending 3 weeks after hospital discharge, patients were encouraged to contact their coordinator (a specially trained nurse or physiotherapist) for continuity of care should they have any health-related concerns. The day after discharge, the coordinator phoned patients in order to evaluate their health status. During the third post-surgery week, patients met with a physiotherapist in an outpatient clinic for a final follow-up on functional performance and general health status.

The E-THAP protocol encouraged patients to access a WIP that incorporates an animated video. The platform was available to patients as soon as they had consented to surgery and participation in the study, which was prior to the information group meeting. Each patient received a login and password, and the research nurse subsequently provided hands-on instructions in how to use the WIP. The login was used to verify patient accessed to the WIP.

The visual approach was designed to reduce the complexity of the information and bridge challenges in health literacy.[28] In line with Internet provision of CCBT,[16] our approach was also

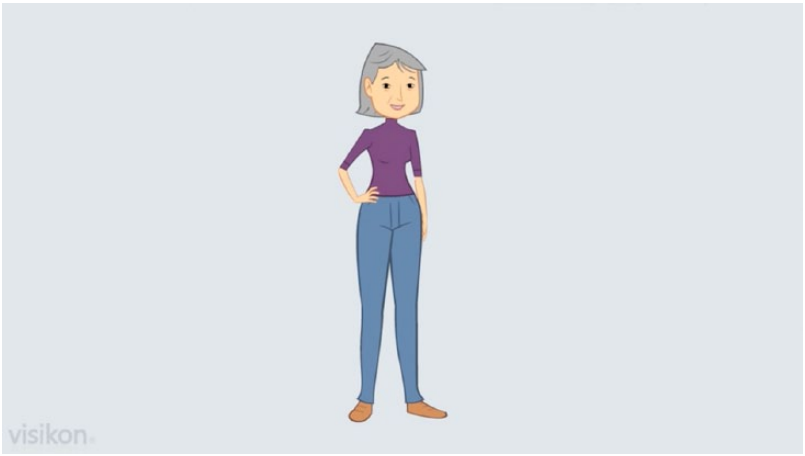

Figure 1. The avatar model (Silkeborg Regional Hospital, Denmark, 2013. 356 × 200 mm (72 × 72 DPI)).

Figure 2. Pre-surgical interaction. Interaction between avatar patient and health providers during anaesthesia (the animation anaesthesia study, Silkeborg Regional Hospital, Denmark, 2013. 356 × 200 mm (72 × 72 DPI)).

designed to reduce the risk of developing catastrophic thoughts and misconceptions in relation to surgery. Our animation comprised a 12-min, two-dimensional (2D) video, comprising 10 sequences to explain the THA procedure, from initial symptoms to post-surgical rehabilitation at home. The animation video showed a patient avatar as a middle-aged woman undergoing the THA procedure, with a male voice-over explaining the procedures (Figures 1 and 2).

The animation displayed on the platform was divided into chapters mirroring the chronology of the THA. The intention was to offer the patient the advantage of knowledge acquisition in familiar surroundings (at home),[29] at a speed dictated by the patient.[30] Research has shown that paced and repeated information appears to improve learning ability both pre- and post-surgically.[20] Furthermore, the platform held written and audio information, in parallel with animation. Animated instructions for physical exercises, bio-feed-back in relation to pain and pain management and

contact information for healthcare professionals in the hospital were provided. Patients could access the platform for up to 3 months after their hospital discharge.

Patient enrolment

The patients receiving standard care, S-THAP (control group), were enrolled consecutively from September 2012 to March 2013. Subsequently, S-THAP was substituted by E-THAP (the intervention group), with patients consecutively recruited to this group from August 2013 to October 2014. The sample size calculation was based on a between-group difference in *Hospital Anxiety and Depression Scale* (HADS)–Anxiety change (primary outcome) from pre-surgical to 3 months post-surgical of 1.5, defined as the minimal clinical relevant difference according to Puhan et al.[31] With an expected standard deviation (SD) of 3.44,[32] a significance level of 0.05 and a power of 0.80, the required sample size was estimated to be 83 patients for each group.

Due to a possible low response rate in some outcome variables, and to accomplish our secondary aim of performing a predictor analysis, we included a total of 299 subjects in the study. The inclusion criterion was a patient undergoing a primary THA surgical procedure; exclusion criteria included cognitive deficiencies and the inability to speak or understand Danish.

Instruments

The HADS[33] is a validated 14-item psychological screening scale that measures symptoms of anxiety and depression (7 items each) with a score range of 0–21. The recommended cut-off in defining symptoms of anxiety or depression is >7.[33]

The *Visual Analogue Scale-Anxiety* (VAS-A)[34] measures self-perceived anxiety from 0 mm (not anxious), to 100 mm (most anxious).[34] VAS-A is validated[35] with a recommended cut-off at >30 mm.

EuroQoL-5 Dimension Questionnaire (*EQ-5D*)-*3L*, including *EQ-VAS*,[36,37] is a 9-item validated instrument that measures health conditions according to the following subscales: mobility, three items; self-care, three items; usual activities, three items; pain/discomfort, three items; and anxiety/depression, three items.[36,37] These subscales are divided into three levels of perceived problem: level 1, no problem; level 2, some problems and level 3, extreme problems. A time trade off (TTO) was used to estimate the EQ-5D index.[36,37] EQ-VAS measures self-perceived health condition from 0 mm (worst) to 100 mm (best).

The *General Self-efficacy Scale* (*GSES*)[38] is a 10-item validated scale to assess optimistic self-belief in coping with a variety of difficult demands in life[38] and is scored between 1 and 4 points, without a cut-off.

Data collection

Data collection pre-surgery and 3 months post-surgery was with a composite questionnaire comprising all instruments (see Figure 3). In addition, VAS-A was measured before admission (prior to the information meeting), just before surgery (at the ward) and again before discharge.

Ethical considerations

The Danish Data Protection Agency[39] provided permission to conduct this study (J. No. 2007-58-0010). Given the absence of any biologic tests, no contribution was required from the Scientific Committee for the County of Central Jutland, or the Biomedical Research Ethics

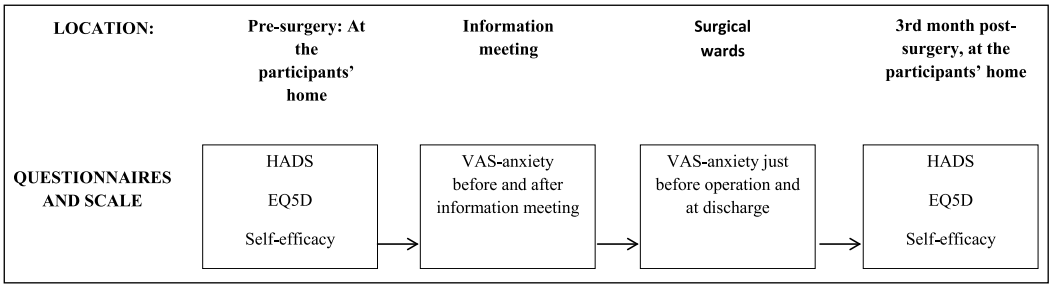

Figure 3. Flowchart of the chronology and location of the data collection process.
HADS: Hospital and Anxiety Scale; EQ-5D: EuroQoL-5 Dimension Questionnaire; VAS: Visual Analogue Scale; WIP: web-based information platform.

Committee. In line with the Helsinki Declaration,[40] patients were informed about the study both verbally and in writing, and all participants gave their written consent to participate.

Statistical analyses and outcomes

Data were double-entered in EpiData 3.1 (Epidata Association, Odense, Denmark), with statistical analyses performed using the STATA 13 software package (StataCorp, College Station, TX, USA). Normally distributed data were described by means and SD, non-normally distributed data, by medians and interquartile range (IQR), with categorical variables described by numbers and proportions. Between-group and within-group differences were tested with an unpaired t test, and a paired t test, or the equivalent non-parametric Wilcoxon rank-sum test. Categorical variables were compared between groups using the chi^2 test or Fisher's exact test for less than five subjects in a category. The primary analysis was a between-group comparison regarding change in HADS score from pre-surgery to 3 months post-surgery. Secondary analyses were (1) a between-group comparison of scores for the remaining outcome variables and (2) a predictive analysis performed for pre-surgical values of VAS-A >30 mm, as the predictive variable (defined as a score >30 mm at a pre-surgery ambulatory visit or information meeting), with scores for outcome variables for VAS-A just before surgery, at hospital discharge and 3 months post-surgery. Due to a substantial proportion of patients not using the platform in the E-THAP group, a subgroup analysis was performed to compare users (those who accessed) versus non-users (those who did not access) of the platform. A log related to each patient's study number on the platform allowed us to the identify users versus non-users.

Results

A total of 299 patients were included: 117 in the S-THAP group and 182 in the E-THAP group (see Figure 4).

No significant differences on baseline characteristics were found between the groups (see Table 1).

The effect of supplementing standard information with web-based assistance

On the primary outcome of HADS, there was no significant improvement for the E-THAP versus the S-THAP group ($p=0.06$). However, all outcomes improved significantly from pre-admission to 3 months post-surgery within groups, but with no between-group differences as illustrated in Table 2.

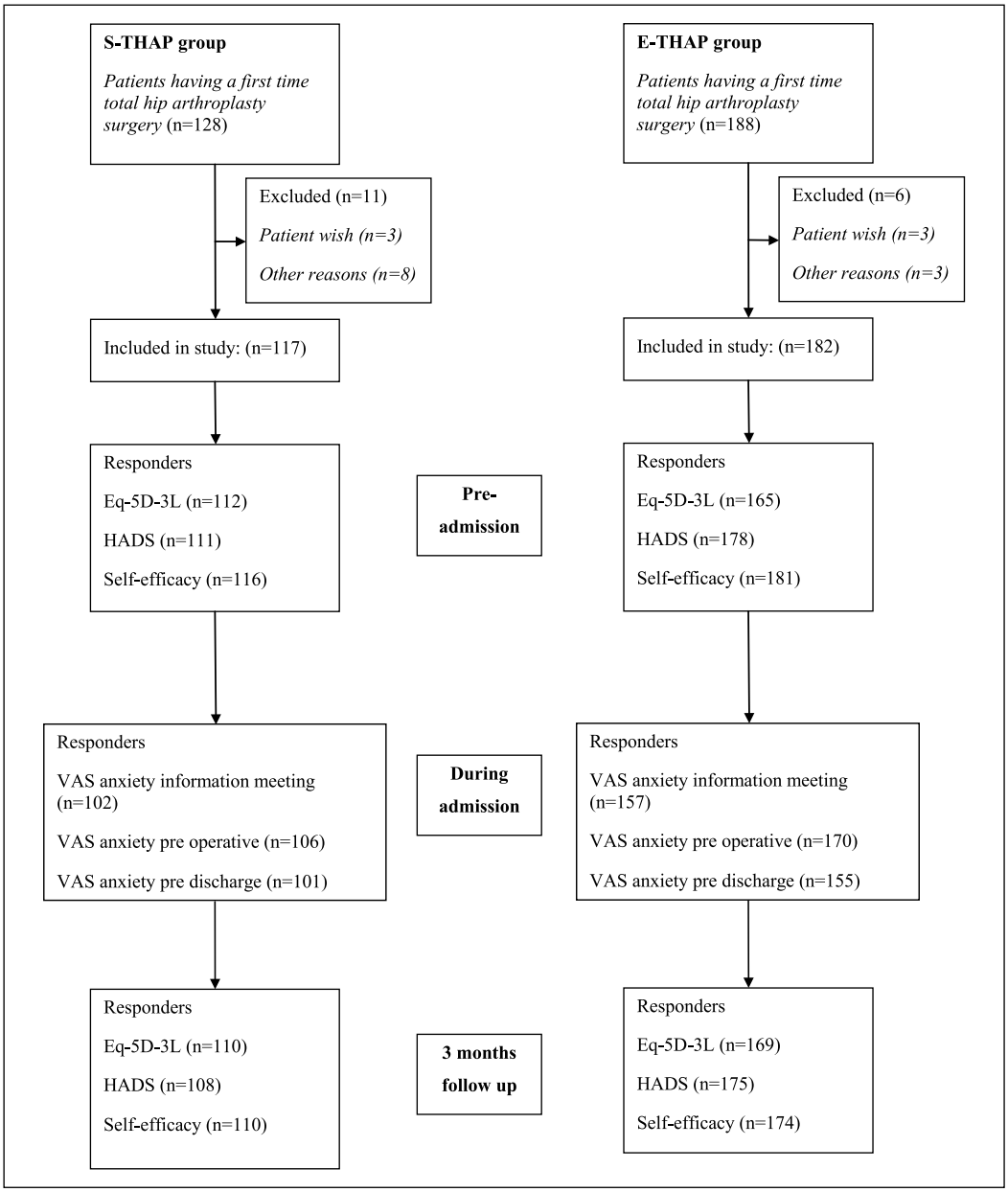

Figure 4. Inclusion and exclusion criteria for participants, with the distribution of responders.

Predictor of psychosocial outcomes within groups

As demonstrated in Tables 3 and 4, pre-surgery anxiety >30 mm VAS was associated with a significantly higher level of anxiety just before surgery for both groups ($p < 0.0001$) and at discharge for the S-THAP group ($p = 0.04$). Patients with pre-surgery anxiety (>30 mm VAS) demonstrated significantly lower self-efficacy after 3 months compared to those with lower pre-surgery anxiety

Table 1. Baseline characteristics for patients with and without access to the web-based information platform.

	Patient demographics (n = 117)	WIP (n = 182)	p Value*
Age (years), mean (SD)	65.7 (10.3)	67.1 (9.1)	0.24 (t test)
Range	36–90	30–85	
BMI (kg/m^2), mean (SD)	27.2 (4.8)	27.7 (5.1)	0.40 (t test)
Female gender, n (%)	61 (52.1)	91 (49.7)	0.68 (chi^2)
Physical status	(n = 116)	(n = 182)	
ASA I, n (%)	48 (41)	58 (32)	
ASA II, n (%)	62 (53)	111 (61)	
ASA III, n (%)	6 (5)	13 (7)	0.23 (chi^2)
Education, n (%)	(n = 105)	(n = 165)	
Low	54 (51)	75 (45)	
Middle	42 (40)	77 (47)	
High	9 (9)	14 (8)	0.56 (chi^2)
Postoperative days in hospital, n (%)			
1	88 (75)	145 (80)	
2	23 (20)	31 (17)	
⩾3	6 (5)	6 (3)	0.59 (chi^2)
EQ-5D	(TTO n = 116, VAS n = 112)	(TTO n = 175, VAS n = 168)	
Health status (TTO), median (IQR)	0.71 (0.63–0.72) (n = 116)	0.66 (0.44–0.72) (n = 175)	0.10[a]
VAS, mean (SD)	65.75 (19.2) (n = 112)	61.2 (20.3) (n = 168)	0.06 (t test)
Self-efficacy, mean (SD)	(n = 116)	(n = 181)	
	30.92 (5.7)	31.78 (5.1)	0.18 (t test)
HADS, n (%)	(n = 110)	(n = 178)	
Anxiety symptoms (>7)	21 (19.1)	32 (18.0)	0.81 (chi^2)
Depression symptoms (>7)	3 (2.7)	10 (5.6)	0.38[b]

WIP: web-based information platform; SD: standard deviation; BMI: body mass index; ASA: American Society of Anaesthesiologists physical status classification – I: healthy patient, II: patient with mild systemic disease, III: patient with severe systemic disease; EQ-5D: EuroQoL-5 Dimension Questionnaire; TTO: time trade off; IQR: interquartile range; VAS: Visual Analogue Scale; HADS: Hospital and Anxiety Scale.
*Between-group comparison: [a]Wilcoxon rank-sum test, [b]Fisher's exact test.

(⩽30 mm); this finding applied to both groups (p = 0.05), with only small differences in median values for S-THAP (Table 3) versus E-THAP (Table 4).

As illustrated in Figure 5, a chronologic plot of the anxiety level 'profile' between groups was comparable.

When comparing the means for all four measurements of VAS anxiety between groups, there were no significant differences between the overall VAS anxiety level for S-THAP versus E-THAP (p = 0.10) (Figure 5).

Subgroup analysis in the E-THAP group for users and non-users of the WIP

In all, 112 of 182 participants in the E-THAP group accessed the platform. As illustrated in Table 5, a subgroup analysis showed that there were neither significant differences between users and non-users at baseline nor changes in outcome.

Table 2. Changes from baseline to 3 months postoperative, values are mean (95% CI).

Variable	Standard information	p value (within group)[a]	WIP	p value (within group)	Difference (between groups)[b]	p value (between groups)[c]
ΔHADS-Anxiety	2.39 (1.7, 3.1) (n = 101)	<0.0001	1.63 (1.2, 2.1) (n = 168)	<0.0001	−0.76 (−0.5, 0.0)	0.06
ΔHADS-Depression	0.90 (0.4, 1.4) (n = 101)	0.0004	1.04 (0.68, 1.39) (n = 168)	<0.0001	0.13 (−0.5, 0.7)	0.65
ΔEQ-5D – VAS	17.7 (14.1, 21.3) (n = 106)	<0.0001	22.4 (19.0, 25.8) (n = 160)	<0.0001	4.74 (−0.2, 9.7)	0.06
ΔEQ-5D – health status	0.23 (0.2, 0.3) (n = 109)	<0.0001	0.27 (0.2, 0.3) (n = 161)	<0.0001	0.05 (−0.0, 0.1)	0.07
ΔSelf-efficacy	1.78 (1.0, 2.6) (n = 109)	<0.0001	1.35 (0.7, 2.0) (n = 173)	0.0001	−0.43 (−1.5, 0.6)	0.43

WIP: web-based information platform; HADS: Hospital and Anxiety Scale; EQ-5D: EuroQoL-5 Dimension Questionnaire; CI: confidence interval.
[a]Paired t test.
[b]Difference in change scores between groups; positive numbers represent larger improvements with WIP than without WIP.
[c]Unpaired t test.

Table 3. VAS anxiety as a predictor for psychosocial outcomes in the group with standard information (phase 1) (n = 111, due to six missings on the predictor variable).

Outcome	n	Preoperative anxiety ⩽30 mm (n = 66) Median (IQR)	n	Preoperative anxiety >30 mm (n = 45) Median (IQR)	p value[a] (between group)
In-hospital					
VAS anxiety *pre-surgery*	60	10.5 (5, 20)	41	38 (26, 68)	<0.0001
VAS anxiety *at discharge*	58	6 (3, 14)	40	13.5 (3, 23.5)	0.04
3-Month follow-up					
EQ-5D – VAS	64	86.5 (77.5, 92)	40	90 (80, 93.5)	0.59
EQ-5D – health status	64	1 (0.8, 1)	40	0.82 (0.76, 1)	0.08
Self-efficacy	64	33.5 (30, 37)	40	33 (28, 36.5)	0.049
HADS,[b] n (%)					
HADS-Anxiety >7	63	3 (4.8)	39	4 (10.3)	0.42[c]
HADS-Depression >7	63	1 (1.6)	39	2 (5.1)	0.56[c]

IQR: interquartile range; VAS: Visual Analogue Scale; EQ-5D: EuroQoL-5 Dimension Questionnaire; HADS: Hospital and Anxiety Scale.
[a]Wilcoxon rank-sum test unless otherwise stated.
[b]HADS, presented as proportion with anxiety/depression symptoms.
[c]Fisher's exact test.

Discussion

This un-controlled, before-and-after study showed no significant psychosocial effects for patients joining E-THAP compared to S-THAP. Any difference between groups in favour of S-THAP was beneath the 1.5 point threshold suggested to indicate minimal clinical relevance. Extensive written preoperative information as used in both S-THAP and E-THAP is a well-known intervention with which to address preoperative anxiety[14] and may explain the lack of any additional benefit when using WIP.

Table 4. VAS anxiety as a predictor for psychosocial outcomes in the group with WIP (phase 3) (n = 173 due to 10 missings in the predictor variable).

Outcome	n	Preoperative anxiety ⩽30 mm (n = 117) Median (IQR)	n	Preoperative anxiety >30 mm (n = 56) Median (IQR)	p value[a] (between group)
In-hospital					
VAS anxiety *pre-surgery*	110	11 (3, 20)	52	38 (20.5, 70.5)	<0.0001
VAS anxiety *at discharge*	98	5 (2, 11)	49	7 (3, 19)	0.23
3-Month follow-up					
EQ-5D – VAS	112	88 (78.5, 95)	52	88.5 (77.5, 91.5)	0.68
EQ-5D – health status	105	1 (0.78, 1)	52	0.82 (0.72, 1)	0.11
Self-efficacy	109	34 (30, 38)	56	33 (29, 36)	0.049
HADS,[b] n (%)					
HADS-Anxiety >7	109	8 (7.3)	55	8 (14.6)	0.14[c]
HADS-Depression >7	109	1 (0.9)	55	3 (5.5)	0.11[d]

IQR: interquartile range; VAS: Visual Analogue Scale; EQ-5D: EuroQoL-5 Dimension Questionnaire; HADS: Hospital and Anxiety Scale.
[a]Wilcoxon rank-sum test unless otherwise stated.
[b]HADS, presented as proportion with anxiety/depression symptoms.
[c]Fisher's exact test.
[d]Chi^2 test.

Regarding the didactic elements of the 'animated information video', contradictory effects on surgery-related anxiety have previously been reported. Kakinuma et al.[24] reported no beneficial effect on pre-surgery anxiety following the provision of animated information. In contrast, a positive effect was identified by Tou et al.[17] Neither study dealt with orthopaedic fast-track surgical programmes, with measurement times also varying between studies. Interestingly, a beneficial effect conferred by the animated video on learning ability was found.[24] When considering the correlate of mental health and learning capacity, improved learning abilities may, at length, help to diminish anxiety. Patient preferences in terms of self-identification with the avatar used in the animation video may also explain the missing psychosocial effect. Based on the learning theory by Bandura,[41] identification with an avatar may have a beneficial effect on learning ability through the theory of vicarious reinforcement.[41] We conducted an ethnographic study on a subset of the total S-THAP group.[42] Fifteen of the most anxious patients (VAS anxiety >30 mm) were included after 3 months of follow-up. Individual preferences were found for two different narrative models, one speaking from the patient's perspective and the other with a more formal instructor perspective (i.e. a healthcare professional), as used in E-THAP. The most anxious patients preferred the narration to reflect the patient's perspective, which contradicts the approach taken in the WIP. Different types of avatar may have also impacted patient cognition, due to their anxiety level. This suggests that individual-specific targeting is appropriate for the didactic components of the WIP, such that these can be aligned with preoperative levels of anxiety.

As all outcomes improved significantly from pre-admission to 3 months after surgery, with no between-group differences, both programmes had positive psychosocial effects. The finding of post-surgical improvement of health status among THA patients is in line with previous findings,[38] and may be related to the surgery in itself,[43] and the extensive preoperative information given.[26] Irrespective of the mode of information delivery, the fast-track concept, with clear expectations of operative course, rehabilitation and patient agency, contributes to the success of THA.

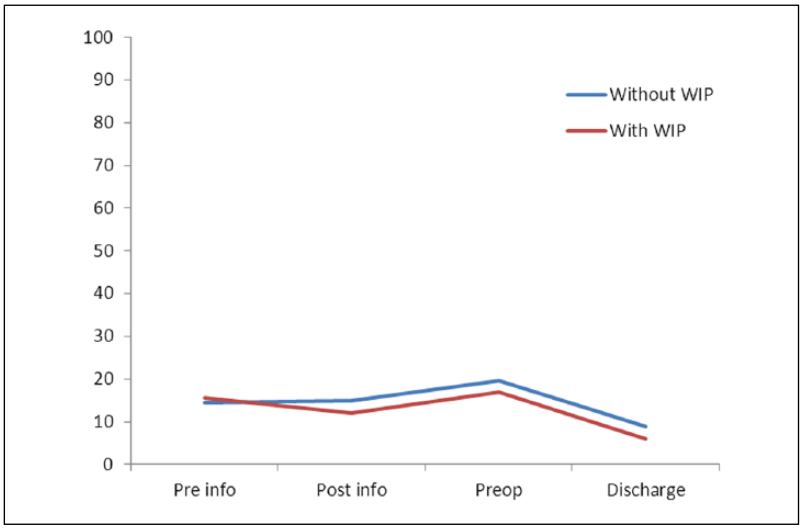

Figure 5. VAS anxiety median scores before and during hospital admission for groups with and without access to the WIP.
HADS: Hospital and Anxiety Scale; EQ-5D: EuroQoL-5 Dimension Questionnaire; VAS: Visual Analogue Scale; WIP: Web-based information platform.

Table 5. Patients in the group with access to WIP, grouped according to its usage.

	Users (n = 112)	Non-users (n = 70)	p Value[a]
HADS-Anxiety symptoms (>7), n/total n (%)			
Baseline	21/111 (19%)	11/67 (16%)	0.67 (chi^2)
3 Months postoperative	9/107 (8%)	8/67 (12%)	0.45 (chi^2)
HADS-Depression symptoms (>7), n/total n (%)			
Baseline	5/111 (5%)	5/67 (7%)	0.41 (chi^2)
3 Months postoperative	3/108 (3%)	2/67 (3%)	0.64 (Fisher's exact)
EQ-5D – health status (TTO)	(n = 109)	(n = 65)	
Baseline, median (IQR)	0.66 (0.5-0.7)	0.66 (0.4–0.8)	
Change baseline–3 months postoperative, mean (SD)	(n = 100)	(n = 61)	
	0.27 (0.2)	0.29 (0.2)	0.53 (t test)
EQ-5D – VAS, mean (SD)	(n = 107)	(n = 58)	
Baseline	62.9 (19.4)	58.0 (22.0)	
Change baseline–3 months postoperative	(n = 99)	(n = 57)	
	20.0 (21.6)	26.0 (21.8)	0.10 (t test)
Self-efficacy, mean (SD)	(n = 112)	(n = 69)	
Baseline	32.3 (4.4)	31.0 (6.0)	
Change baseline–3 months postoperative	(n = 107)	(n = 66)	
	1.0 (4.6)	1.9 (4.5)	0.24 (t test)

WIP: web-based information platform; HADS: Hospital and Anxiety Scale; EQ-5D: EuroQoL-5 Dimension Questionnaire; TTO: time trade off; IQR: Interquartile range; SD: standard deviation, VAS: Visual Analogue Scale.
[a]Test of difference between users and non-users.

Having preoperative anxiety >30 mm VAS was associated with significantly higher anxiety levels during admission in both groups, significantly lower self-efficacy in postoperative tasks after 3 months for both groups and a significantly higher anxiety level at discharge in the S-THAP group. Previous studies have shown that anxiety measured in the anxiety/depression dimension of EQ-5D is a strong predictor of pain and satisfaction after total hip replacement surgery.[44] Additionally, anxiety is a predictor for pain, opioid consumption and level of function pre- and post-surgery.[14,44] Identifying those patients who are most preoperatively anxious is therefore of clinical relevance in a resource-demanding fast-track setting, for which success is contingent on a sufficient level of mental and educational patient resources. The validity and reliability of using the VAS anxiety score to determine pre-surgery anxiety, and as a predictor for a patient's mental state while participating in a fast-track programme, must be further studied.

A subgroup analysis showed no significant differences between users and non-users at baseline and in terms of outcome. However, a trend was evident with respect to a larger positive change in HADS-anxiety for users compared to non-users. Conversely, non-users seemed to have slightly larger improvements for the remaining outcomes, suggesting no overall positive effect from use of the WIP. Providing a pamphlet with extensive information for both groups may have had the effect of influencing some participants not to access the WIP. However, different perceptions of usefulness (i.e. utilitarian motivation) may have influenced the choice of whether to access the WIP, as reported for users and non-users of a mobile application.[45] More information on individual preferences are needed[45] before any firm conclusions can be drawn; these may contribute to our ability to predict how and why groups chose to engage with different modes of health informatics.

Despite no psychosocial effect of the intervention, animated information is still used in the department as former THA patients with a VAS-A >30 mm find the information trustworthy and the displayed THAP recognisable.[42] The department is continuing to study the effects of WIP for patients undergoing THAP and other surgical treatments.

Study limitations

Compared to a randomised controlled design, this un-controlled design may have compromised data quality given its suboptimal ability to measure the effect of a given intervention. However, the choice of this study design was predicated by economic and logistic considerations, as a randomised study would have required a greater staff allocation, with duplicated in-parallel information meetings, and separate wards.

The un-controlled design also introduces a risk of bias due to the unequal distribution of confounding factors between groups. However, these quality flaws may be compensated for by the consecutive recruitment as well as baseline characteristics, which indicated equal distributions, at least among the measured variables, and support the validity of our finding of no significant effects of the E-THAP programme on psychosocial outcomes. The size of the study, which complied with a sample size calculation, also strengthens its validity. However, missing data, as indicated in Figure 4, is a weakness.

In this study, multiple comparisons were made which increases the risk of type I errors. Consequently, the significant differences that we found should be interpreted with caution and in relation to the clinical relevance of the findings. All questionnaires used in the study are validated and commonly used in studies for patients undergoing THA.[32,44,46]

Conclusion

This study documented improved psychosocial effects for fast-track THAP, with no additional effect of web-based and animated information. Anxiety was a possible predictor of a higher

anxiety level during hospitalisation and lower self-efficacy post-surgery. Using VAS-A to identify the most nervous patients must now be tested in a further validation study. Likewise, an evaluation of the psychosocial effect of E-THAP should be conducted in a randomised controlled trial, taking individual coping preferences as well as avatar preferences into account. Identifying utilitarian motivational factors may be necessary to improve our understanding of the psychosocial benefits of E-THAP and allow us to target fast-track THAPs that incorporate psychosocial challenges.

Declaration of conflicting interests

The author(s) declared the following potential conflicts of interest with respect to the research, authorship, and/or publication of this article: Co-author Martin Vesterby is a co-founder and partner of Visikon Inc. that produced the animation video used in the study.

Funding

The author(s) received no financial support for the research, authorship and/or publication of this article.

References

1. Norlyk A and Harder I. After colonic surgery: the lived experience of participating in a fast-track programme. *Int J Qual Stud Health Well-Being* 2009; 4: 170–180.
2. Norlyk A and Harder I. Recovering at home: participating in a fast-track colon cancer surgery programme. *Nurs Inq* 2010; 18: 165–173.
3. Woolf AD and Pfleger B. Burden of major musculoskeletal conditions. *Bull World Health Organ* 2003; 81: 646–656.
4. Changulani M, Kalairajah Y, Peel T, et al. The relationship between obesity and the age at which hip and knee replacement is undertaken. *J Bone Joint Surg Br* 2008; 90: 360–363.
5. Wilmore D and Kehlet H. Management of patients in fast track surgery. *BMJ* 2001; 322: 473–476.
6. Wilmore D. From Cuthbertson to fast-track surgery: 70 years of progress in reducing stress in surgical patients. *Ann Surg* 2002; 236: 643–648.
7. Kehlet H and Wilmore D. Evidence-based surgical care and the evolution of fast-track surgery. *Ann Surg* 2008; 248: 189–198.
8. Husted H, Solgaard S, Hansen T, et al. Care principles at four fast-track arthroplasty departments in Denmark. *Dan Med Bull* 2010; 57: A4166.
9. Ibrahim M, Khan M, Nizam I, et al. Peri-operative interventions producing better functional outcomes and enhanced recovery following total hip and knee arthroplasty: an evidence-based review. *BMC Med* 2013; 11: 37.
10. Kehlet H and Søballe K. Fast-track hip and knee replacement – what are the issues? *Acta Orthop* 2010; 81: 271–272.
11. Ayers D, Franklin P and Ring D. The role of emotional health in functional outcomes after orthopaedic surgery: extending the biopsychosocial model to orthopaedics: AOA critical issues. *J Bone Joint Surg Am* 2013; 95: e165(1–7).
12. Rosenfeld R. Anxiety and learning. *Teach Sociol* 1978; 5: 151–166.
13. Frederiks S, Sidani S and Shugurensky D. The effect of anxiety on learning outcomes post-CABG. *Can J Nurs Res* 2008; 40: 127–140.
14. Sjöling M, Nordahl G, Olofsson N, et al. The impact of preoperative information on state anxiety, postoperative pain and satisfaction with pain management. *Patient Educ Couns* 2003; 51: 169–176.
15. McDonald S, Page M, Beringer K, et al. Preoperative education for hip or knee replacement. *Cochrane Database Syst Rev* 2014; 5: CD003526.
16. Kalthenthaler E, Brazier J, De Nigris E, et al. Computerised cognitive behaviour therapy for depression and anxiety update: a systematic review and economic evaluation. *Health Technol Assess* 2006; 10: iii, xi-xiv, 1–168.

17. Tou S, Tou W, Mah D, et al. Effect of preoperative two-dimensional animation information on perioperative anxiety and knowledge retention in patients undergoing bowel surgery: a randomized pilot study. *Colorectal Dis* 2013; 15: e256–e265.
18. Delp C and Jones J. Communicating information to patients: the use of cartoon illustrations to improve comprehension of instructions. *Acad Emerg Med* 1996; 3: 264–270.
19. Houts S, Doak C, Doak L, et al. The role of pictures in improving health communication: a review of research on attention, comprehension, recall, and adherence. *Patient Educ Couns* 2006; 61: 173–190.
20. Leiner M, Handal G and Williams D. Patient communication: a multidisciplinary approach using animated cartoons. *Health Educ Res* 2004; 19: 591–595.
21. Calderon JL, Shaheen M, Hays RD, et al. Improving diabetes health literacy by animation. *Diabetes Educ* 2014; 40: 361–372.
22. Heikkinen K, Salanterä S and Leino-Kilpi H. Ambulatory orthopaedic surgery patients' knowledge with internet-based education. *Methods Inf Med* 2012; 51(4): 295–300.
23. Swearingen C, McCollum L, Daltroy L, et al. Screening for low literacy in a rheumatology setting: more than 10% of patients cannot read 'cartilage', 'diagnosis', 'rheumatologist', or 'symptom'. *J Clin Rheumatol* 2010; 16: 359–364.
24. Kakinuma A, Nagatani H, Otake H, et al. The effects of short interactive animation video information on preanesthetic anxiety, knowledge, and interview time: a randomized controlled trial. *Anesth Analg* 2011; 112: 1314–1318.
25. Johansson K, Nuutila L, Virtanen H, et al. Preoperative education for orthopaedic patients: systematic review. *J Adv Nurs* 2005; 50: 212–223.
26. Pellino T, Tluczek A, Collins M, et al. Increasing self-efficacy through empowerment: preoperative education for orthopaedic patients. *Orthop Nurs* 1998; 17: 48–51, 54–59.
27. Learmonth I, Young C and Rorabeck C. The operation of the century: total hip replacement. *Lancet* 2007; 370(9597): 1508–1519.
28. Hoppe D, Denkers M, Hoppe F, et al. The use of video before arthroscopic shoulder surgery to enhance patient recall and satisfaction: a randomized-controlled study. *J Shoulder Elbow Surg* 2014; 23: e134–e139.
29. Ong J, Miller P, Appleby R, et al. Effect of a preoperative instructional digital video disc on patient knowledge and preparedness for engaging in postoperative care activities. *Nurs Clin North Am* 2009; 44: 103–115, xii.
30. Adams M, Chen B, Mokowski R, et al. Multimedia approach to preoperative adenotonsillectomy counseling. *Otolaryngol Head Neck Surg* 2012; 146: 461–466.
31. Puhan M, Frey M, Büchi S, et al. The minimal important difference of the Hospital Anxiety and Depression Scale in patients with chronic obstructive pulmonary disease. *Health Qual Life Outcomes* 2008; 6: 46.
32. Mikkelsen L, Petersen M, Søballe K, et al. Does reduced movement restrictions and use of assistive devices affect rehabilitation outcome after total hip replacement? A non-randomized, controlled study. *Eur J Phys Rehabil Med* 2014; 50: 383–393.
33. Herman C. International experiences with the Hospital Anxiety and Depression Scale. A Review of validation data on clinical results. *J Psychosom Res* 1997; 42: 17–41.
34. Youngblut J and Casper G. *Focus on psychometrics*: Single-Item indicators in nursing research. *Res Nurs Health* 1993; 16: 459–465.
35. Facco E, Stillini E, Bacci C, et al. Validation of visual analogue scale for anxiety (VAS-A) in preanesthesia evaluation. *Minerva Anestesiol* 2013; 79: 1389–1395.
36. EuroQol Group. EuroQol – a new facility for the measurement of health-related quality of life. *Health Policy* 1990; 16: 199–208.
37. Brooks R. EuroQol: the current state of play. *Health Policy* 1996; 37: 53–72.
38. Schwarzer R and Jerusalem M. General self-efficacy scale. In: Weinman S and Johnston M (eds) *Measures in health psychology: a user's portfolio. Causal and control beliefs*. Windsor: NFER-NELSON, 1995, pp. 35–37.
39. Danish Regions. Health, http://www.regioner.dk/sundhed/sundheds-it (accessed 17 June 2013).

40. *World Medical Association (WMA) Declaration of Helsinki: ethical principles for medical research involving human subjects* (52nd World Medical Association general assembly). Edinburgh, http://www.wma.net/en/30publications/10policies/b3/
41. Bandura A. *Social learning theory*. Englewood Cliffs, NJ: Prentice Hall, 1977.
42. Høybye M, Vesterby M and Jørgensen L. Producing patient-avatar identification in animation video information on spinal anesthesia by different narrative strategies. *Health Informatics J*. Epub ahead of print 23 December 2014. DOI: 10.1177/1460458214560636.
43. Sharma W, Morgan P and Cheng E. Factors influencing early rehabilitation after THA: a systematic review. *Clin Orthop Relat Res* 2009; 467: 1400–1411.
44. Rolfson O, Le Dahlberg Nilsson J, et al. Variables determining outcome in total hip replacement surgery. *J Bone Joint Surg Br* 2009; 91: 157–161.
45. Verkasalo H, López-Nicolás C, Molina-Castillo F, et al. Analysis of users and non-users of smartphone applications. *Telemat Inform* 2010; 27: 242–255.
46. Mahomed N, Liang M, Cook M, et al. The importance of patient expectations in predicting functional outcomes after total joint arthroplasty. *J Rheumatol* 2002; 29: 1273–1279.

Health Informatics Journal

Article

Extracting and analyzing ejection fraction values from electronic echocardiography reports in a large health maintenance organization

Health Informatics Journal
2017, Vol. 23(4) 319–328

DOI: 10.1177/1460458216651917
journals.sagepub.com/home/jhi
SAGE

Fagen Xie and Chengyi Zheng
Kaiser Permanente Southern California, USA

Albert Yuh-Jer Shen
Kaiser Permanente Los Angeles Medical Center, USA

Wansu Chen
Kaiser Permanente Southern California, USA

Abstract

The left ventricular ejection fraction value is an important prognostic indicator of cardiovascular outcomes including morbidity and mortality and is often used clinically to indicate severity of heart disease. However, it is usually reported in free-text echocardiography reports. We developed and validated a computerized algorithm to extract ejection fraction values from echocardiography reports and applied the algorithm to a large volume of unstructured echocardiography reports between 1995 and 2011 in a large health maintenance organization. A total of 621,856 echocardiography reports with a description of ejection fraction values or systolic functions were identified, of which 70 percent contained numeric ejection fraction values and the rest (30%) were text descriptions explicitly indicating the systolic left ventricular function. The 12.1 percent (16.0% for male and 8.4% for female) of these extracted ejection fraction values are <45 percent. Validation conducted based on a random sample of 200 reports yielded 95.0 percent sensitivity and 96.9 percent positive predictive value.

Keywords

Echocardiography reports, ejection fraction, information retrieval, left ventricle, natural language processing

Corresponding author:
Fagen Xie, Department of Research & Evaluation, Kaiser Permanente Southern California, 100 South Los Robles Avenue, 2nd Floor, Pasadena, CA 91101, USA.
Email: fagen.xie@kp.org

Introduction

Electronic medical record (EMR) systems, mandated by the federal government, are being increasingly used to improve care quality and patient safety.[1] EMR systems often capture rich clinical information such as signs, symptoms, severity, and disease onset, many of which are only available in unstructured data. A number of natural language processing (NLP) systems have been successfully developed in various clinical domains with varying focuses.[2–9] For example, the Clinical Text Analysis and Knowledge Extraction System handles general clinical documents.[2] It processes clinical notes to identify and annotate various clinical named entities including drugs, diseases/disorders, signs/symptoms, anatomical sites, and procedures. The Cancer Text Information Extraction System identifies potential cancer cases based on pathology reports.[3] The Unified Medical Language System[4] provides a mapping structure among the vocabularies in biomedical sciences and also provides a comprehensive thesaurus and ontology of biomedical concepts to facilitate NLP. The Medical Language Extraction and Encoding System was initially designed to process the radiology reports, and then subsequently extended to other biomedical domains such as mammography reports and discharge summaries.[5] Other examples include the application peFinder for document-level classification of computed tomography (CT) pulmonary angiography reports,[6] the NegEx for identifying negated findings,[7] and the applications for radiology reports and other clinical notes.[8,9] These systems provide good frameworks for the biomedical research community to process various unstructured text data.

Left ventricular ejection fraction (LVEF) is an important diagnostic measure of overall cardiac health and is one of the most powerful prognostic indicators in patients with cardiovascular diseases.[10–18] An ejection fraction (EF) measure represents the percentage of blood ejected from the left ventricle with each systole. Healthy individuals typically have EF measures between 50 and 70 percent. An abnormal EF may be an early or late manifestation of cardiac diseases, but its presence invariably portends an unfavorable prognosis, regardless of whether the underlying cause is coronary, valvular, infectious, inflammatory, systemic (e.g. hypertension), idiopathic, or any other etiology. A reduced EF indicates impaired systolic function, which is likely to occur after myocardial infarction, obstructive or regurgitant valvular heart disease, myocarditis, chronic hypertension, and many other conditions.[10–14] EF values <45 percent are associated with higher cardiovascular risk and mortality.[13,14] In many cases, an improvement of EF through therapy is associated with improvements in symptoms and prognosis. Practice guidelines include EF in the treatment decision tree for many conditions.[15,16] For example, in patients with asymptomatic aortic or mitral regurgitation, the patient's EF is a major determinant of the timing of surgical valve repair or prosthetic valve replacement. Beyond its immediate clinical utility, the EF is central to many regulatory agencies' decision for reimbursement or quality of care assessment. For instance, the use of implantable cardioverter-defibrillators for the primary prevention of sudden cardiac death is reimbursable by the Center for Medicare and Medicaid Services only for patients with EFs below a certain value (usually 30%–35%, depending on the substrate).[17] Because of the importance of EF in many cardiac conditions, a wide range of prospective clinical trials and retrospective cohort studies involving cardiac maladies list EF values among the basic and essential characteristics of the cohort.[18,19] Although several other imaging modalities may be used to assess LVEF, including radiocontrast LV angiogram, radionuclide gated-averaged or first-pass angiogram, and magnetic resonance imaging, transthoracic echocardiography (TTE) accounts for the vast majority of clinically measured EF values.[20]

The digital records of individual patient studies have been available to practitioners for over a decade within Kaiser Permanente Southern California (KPSC) medical system, long before the full-blown EMR system was implemented in mid-2000s. However, there have been only a few

registry-type structured databases that were collected for patient care purposes that may have included EF as part of the "case-report form." Majority of EF values can only be found in unstructured echocardiography (ECHO) reports. If one needs to identify patients who had a myocardial infarction in the last 5 years whose EF was 30 percent or less, manual review of all the patients' transthoracic echocardiographic reports is necessary. A systematic and reliable method of extracting EF values of large cohorts or population has major beneficial implications and utilities. Recently, an NLP application called CUIMANDREef was developed to classify patients with EF <40 versus >40 percent based on a sample of 765 ECHO reports of seven Veteran Affairs (VA) medical centers.[21] The authors developed a set of regular expression rules to extract EF concept and value from the ECHO reports and assigned a score to determine the classification of EF. (A patient whose ECHO report that was assigned a score > 0 was classified as EF >40% while as a patient whose report with a score <0 indicated EF <40%.) The actual values of EF were not reported by the CUIMANDREef. The purpose of this study is to develop a computerized NLP algorithm to extract the numeric EF values or text descriptions from the unstructured ECHO reports in a more general and diversity population within a large healthcare maintenance organization. We will also validate the performance of the algorithm and report the distribution of EF values by patient characteristics.

Methods

KPSC is an integrated large health maintenance organization (HMO) that provides services to over 4.0 million members in 14 hospitals and 214 medical offices located in 10 counties of Southern California. Members are insured under employer-sponsored plans, individually purchased plans, and Medicare or state-subsidized programs for the indigent.[22] Members' medical and demographic information is captured in a complete EMR system containing the free-text clinical notes (e.g. ECHO reports, hospital discharge summaries, and outpatient progress notes). This study included the ECHO reports stored in the EMR system that dated back to 1995. The study was approved by our Institutional Review Board.

Types of ECHO reports

The electronic ECHO reports were extracted from the data repository of our EMR system for the period between 1 January 1995 and 31 December 2011. The predominant types of ECHO performed and documented within the EMR system were standard two-dimensional TTE, TTE with treadmill or dobutamine stress, transesophageal echocardiography (TEE), and fetal ECHO. The standard measure of EF is with the modified Simpson's rule, also known as the stacked-disc method. During the acquisition of the images, the sonographer traces the LV endocardial borders in two planes (biplane) in diastole and again in systole. The difference between the two phases yields the EF. Although visually estimated EF values may be reported for all these studies, the protocol to ascertain EF is consistently used only for TTE studies. Therefore, we excluded the following types of ECHO. First, we excluded TEE reports, because the EF values obtained from transesophageal views are not directly comparable to those of standard transthoracic method and protocol. Furthermore, TEEs do not typically show the entire LV so global EF cannot be accurately estimated. Second, stress ECHOs, including dobutamine stress and treadmill stress testing, were excluded because they focus on the detection of subtle wall motion abnormalities and use a different protocol. LV function is estimated visually rather than using the modified Simpson's method. The numerical EF values are not routinely reported with stress ECHOs. Finally, fetal ECHOs were removed because they are used to evaluate the baby's heart for problems during the second

trimester of pregnancy. Other uncommon cardiac studies using ultrasound such as intracardiac ECHO[23] and intravascular ultrasound[24] do not assess EF, and therefore were purposely excluded.

EF extraction algorithm

The EF values in ECHO reports are not uniformly written. Extracting and formatting the data presents big challenges due to the following reasons. First, the EF function may be reported as a numeric value (for example, 35% or 55%) or described by a text string such as "normal," "reduced," or "severe." Second, variations exist in the documentation of EF numeric values or text descriptions associated with the keywords or phrases of our interest (e.g. EF, LVEF, cardiac function). For example, EF values may be reported as "EF was only 30%," "EF = 24%," "LVEF: 55%," or "estimated EF of 40%." Third, some of the EF numeric values or text descriptions that are mentioned in the study ECHO reports refer to historical ECHO reports either inside or outside of our health system. These values need to be removed because they do not represent the actual EF values at the time the ECHO was performed. Finally, there are occasions in which a report contained conflict information on the EF values or text descriptions indicating systolic LV functions. The process to extract the EF values/text descriptions from the ECHO reports is listed below and also shown in Figure 1.

Step 1: segment the reports. This step segmented the ECHO reports into sections, paragraphs, and then to sentences.

Step 2: retrieve EF values or text descriptions. A list of predefined EF keywords or phrases (see Appendix 1 for details) was identified based on ECHO practice guidance[25] and knowledge of conventional usage of an experienced clinical cardiologist. For a specific report, an EF value or text description could appear either prior to or following the predefined keywords or phrases, and thus the searching algorithm could be tedious. First, the keywords or phrases were searched within each segmented sentence in the report. If one of the predefined keywords was identified in a sentence, the numeric values associated with the keyword would be first searched backward and forward within the same sentence starting from the position where the predefined keyword was found. If a numeric value was not found, the algorithm started to search for predefined text description such as "normal," "poor," or "severe" (see Appendix 2(1) and (2)) associated with the keyword. The EF value or text description closest to the predefined keyword was selected if multiple values or text descriptions were found.

Step 3: exclude historic and negated EF values or text descriptions. As mentioned earlier, some ECHO reports include patients' previous EF values or text descriptions indicating systolic LV functions. Therefore, efforts were made to exclude historical results. This was performed by searching each sentence to examine whether or not historical phrases such as "previous" and "history" (see Appendix 2(3)) were associated with the defined keywords listed in Appendix 1. A simple negation algorithm was also applied to each sentence to rule out the negation associated with these EF keywords (see Appendix 2(4)).

Step 4: finalize the EF value for each report. If multiple EF values or text descriptions were found in a report, the final value is determined using the following rules. The values or text descriptions that explicitly describe EF or LV function had the highest priority, followed by those of generic ventricular or cardiac function or contractility. The values or text descriptions pertaining to dysfunction had the lowest priority (see Appendix 1 for details). If multiple EF values or text descriptions had the same priority, the one indicting the worse EF function was selected. For example, if both "moderate"

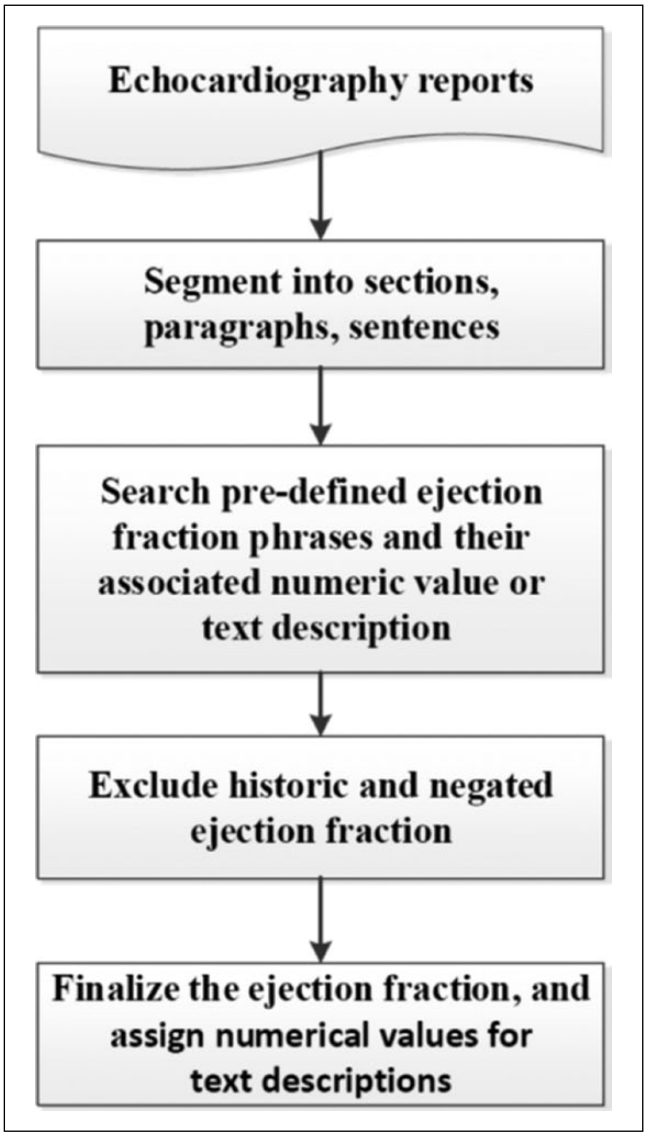

Figure 1. Ejection fraction value retrieval process.

and "severe" descriptions were found to be associated with a keyword listed in Appendix 1, the "severe" one was selected. Finally, the text descriptions of EF were transformed into numerical EF values (see Appendix 3 for details) according to the cardiologist' guidance to facilitate data analysis or reporting. If the original reported EF value was a range, the mid-point of the range was used.

Validation sample

A sample of 200 ECHO reports were randomly selected from the entire ECHO reports included in the study. An experienced clinical cardiologist reviewed each report to abstract the EF value or the

text description indicating systolic LV functions and documented the specific comments if the EF description was not available. The results served as the "gold standard" to be compared with the results from the computerized algorithm. The measures of accuracy, including sensitivity and positive predictive value (PPV), were calculated. Sensitivity was defined as the number of reports in which EF values or text descriptions were *correctly* extracted by the computerized algorithm divided by the total number of reports in which EF values or text descriptions were retrieved by the cardiologist. PPV was the defined as the number of reports in which EF values or text descriptions were *correctly* extracted by the computerized algorithm divided by the total of number of EF values or text descriptions extracted by the computerized algorithm.

Results

The EF numeric values or text descriptions were successfully retrieved from a total of 621,856 ECHO reports between 1995 and 2011 via the developed computerized algorithm. Of these EF results, nearly 70 percent were numeric values while the rest were text descriptions. After these text descriptions were converted into the numerical values using the rules specified in Appendix 3, the distribution of EF values by age and gender is shown in Table 1. The overall number of subjects with EF values was similar between male (49.5%) and female (50.5%). The distribution of EF values among gender was significantly different ($p < 0.01$). Compared to females, males were more likely to have EF values <45 percent (16.0% for male and 8.4% for female). The distribution of EF values by age group was also significantly different ($p < 0.01$). Patients aged 45 years or older had higher percentage with EF values <45 percent compared to patients aged younger than 45 years (13.0% versus 7.5%).

Table 2 shows the comparison of the computerized results versus the manual results retrieved from the 200 randomly selected ECHO reports. The computerized algorithm achieved very high accuracy (sensitivity = 95.0% and PPV = 96.9%).

Discussions

TTEs are among the most commonly obtained imaging studies for hospitalized patients. The primary indication for TTE is to assess LV function, specifically LVEF. Most patients suspected of having a significant cardiac pathology will have undergone a transthoracic echocardiographic study.[20] In many large cohorts of cardiac patients, a significant proportion would have had a TTE in the past. But these values usually were reported in the free-text ECHO reports and unable to be used directly. We developed and implemented a computerized text processing algorithm to extract EF (numerical value or text description if the numerical value was not available) from the unstructured ECHO reports in a large HMO. Compared to the manual chart review results, these computerized extracted EF values or text descriptions achieved a high level of performance (sensitivity (recall) 95.0% and PPV (precision) 96.9%), which were comparable to the sensitivity of 88.9 percent and PPV of 95.0 percent at the concept-level classification of EF of <40 percent in the study conducted at VA.[21] The small percentage (~5%) error rate (either false-positive or false-negative) in our study was caused by either failure to select the correct EF value when multiple EF values appeared in the same report or the unsuccessful efforts to pick up the EF value when the text descriptions of EF mingled with text for multiple things. For example, in the sentence "The LV and RV are normal in size and function," "LV" and "function" are separated by other words, and therefore was not identified as a predefined phrase, although the word "normal" is part of the sentence.

Table 1. The distribution of retrieved ejection fraction values by gender and age group.

	Ejection fraction values (N = 621,856)						
	<15%	15%–29%	30%–36%	37%–44%	45%–55%	>55%	All
Gender							
Male	2081 (0.68%)	17,640 (5.73%)	13,616 (4.43%)	15,954 (5.19%)	127,686 (41.50%)	130,685 (42.48%)	307,662
Female	1091 (0.35%)	8879 (2.83%)	7436 (2.37%)	8792 (2.80%)	122,306 (38.93%)	165,690 (52.73%)	314,194
Age group (years)							
<18	36 (0.28%)	263 (2.04%)	1037 (8.04%)	336 (2.60%)	6741 (52.26%)	4486 (34.78%)	12,899
18–24	36 (0.28%)	208 (1.60%)	143 (1.10%)	172 (1.32%)	5132 (39.41%)	7331 (56.30%)	13,022
25–44	297 (0.44%)	1884 (2.80%)	1155 (1.72%)	1386 (2.06%)	26,791 (39.83%)	35,751 (53.15%)	67,264
45–64	1114 (0.53%)	9024 (4.29%)	6377 (3.04%)	7682 (3.66%)	84,451 (40.19%)	101,462 (48.29%)	210,110
65–75	808 (0.55%)	6753 (4.64%)	5543 (3.80%)	6586 (4.52%)	59,489 (40.84%)	66,501 (45.65%)	145,680
75+	881 (0.51%)	8387 (4.85%)	6797 (3.93%)	8584 (4.97%)	67,388 (38.98%)	80,844 (46.76%)	172,881

Table 2. Performance of the natural language processing algorithm for ejection fraction extraction.

	Confirmed by cardiologist		
	Same	Different	All
Computerized results			
With EF value/text description	190	6	196
Without EF value/text description	0	4	4
Performance			
Sensitivity	95.0%		
Positive predictive value (PPV)	96.9%		

EF: ejection fraction.
The category was defined as "Same" if the computerized result and manual result were identical; otherwise, it was defined as "Different." Sensitivity = number of EF correctly extracted by the computerized algorithm/number of EF retrieved by cardiologist. PPV = number of EF correctly extracted by the computerized algorithm/total of number of EF extracted by computerized algorithm.

To our knowledge, the vast majority of ECHO reports anywhere in the United States are employed by either case report forms, whereby the values are filled in predefined boxes such as LV size, EF, mitral gradient, aortic gradient and pulmonary artery pressure, or dictated reports (the ones being used in this study). For case report forms, it can easily locate the EF value as long as they are reported while there has not been any systematic and automatic way to retrieve the EF value from the dictated report other than reading the report. Therefore, the computerized algorithm developed in our study should be able to apply for any ECHO reports that are narrated and transcribed although the development was based on ECHO reports within our healthcare organization.

The computational time to process ECHO reports depends on the software and hardware configurations such as number and processing power of central processing unit (CPU) and memory. In our study, the algorithm was implemented through Perl scripts and processed on a high-performance Unix server. With CPU of 2.9 GHz and 128G RAM, 70k ECHO reports were processed within 1 h by a single processing job. Our algorithm can also be easily implemented by other programming languages such as Python. With minor modification, the algorithm can also be integrated into other note processing tools.

We acknowledge several potential limitations of this study. First, integration of our study computerized algorithm into the other tools or implementation of this computerized algorithm in other settings may yield some variable results due to the variation in format and presentation of clinical reports, but the accuracy should not be essentially different because the keywords or concepts used for EF extraction are not specified or limited to any fixed/strict formatted ECHO reports. Second, compared to the NegEx algorithm,[7] our study applied a simple rule of the description of negation and the description of history terms to exclude both negation and history conditions in ECHO reports. Despite these limitations, our study developed a computerized algorithm to retrieve EF values from the ECHO reports in a systematic and automated way. This computerized algorithm produced high accuracy and can provide great potential values for improving patient care managements, such as congestive heart failure and other implications. Our algorithm can serve as a starting point for the creation of a more general or global approach. Integration of the algorithm to general NLP tools could yield a more robust and reliable algorithm. Additionally, the incorporation of the knowledge of health professionals allows for a more sophisticated methodology.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

References

1. Committee on Quality of Health Care in America, Institute of Medicine. *Crossing the quality chasm: a new health system for the 21st century*. Washington, DC: Institute of Medicine, 2001.
2. Savova GK, Masanz JJ, Ogren PV, et al. Mayo clinical text analysis and knowledge extraction system (cTAKES): architecture, component, evaluation and applications. *J Am Med Inform Assoc* 2010; 17: 507–513.
3. Crowley RS, Castine M, Mitchell K, et al. caTIES: a grid based system for coding and retrieval of surgical pathology reports and tissue specimens in support of translational research. *J Am Med Inform Assoc* 2010; 17: 253–264.
4. The Unified Medical Language System, http://www.nlm.nih.gov/research/umls (accessed 6 October 2015).
5. Friedman C. A broad-coverage natural language processing system. *Proc AMIA Symp* 2000: 270–274.
6. Chapman BE, Lee S, Kang HP, et al. Document-level classification of CT pulmonary angiography reports based on an extension of the context algorithm. *J Biomed Inform* 2011; 44: 728–737.
7. Chapman WW, Bridewell W, Hanbury P, et al. A simple algorithm for identifying negated findings and diseases in discharge summaries. *J Biomed Inform* 2001; 34: 301–310.
8. Friedman C, Alderson PO, Austin J, et al. A general natural-language text processor for clinical radiology. *J Am Med Inform Assoc* 1994; 1: 161–174.
9. Friedman C, Shagina L, Lussier Y, et al. Automated encoding of clinical documents based on natural language processing. *J Am Med Inform Assoc* 2004; 11: 392–402.
10. Rosamond W, Flegal K, Furie K, et al. Heart disease and stroke statistics—2008 update: a report from the American Heart Association Statistics Committee and Stroke Statistics Subcommittee. *Circulation* 2008; 117: e25–e146.
11. Jessup M, Abraham W, Casey D, et al. 2009 focused update: ACCF/AHA Guidelines for the Diagnosis and Management of Heart Failure in Adults—a report of the American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines. *J Am Coll Cardiol* 2009; 53: 1343–1382.
12. Little WC. Hypertension, heart failure, and ejection fraction. *Circulation* 2008; 118: 2223–2224.
13. Solomon SD, Anevekar N, Skali H, et al. Influence of ejection fraction on cardiovascular outcomes in a broad spectrum of heart failure patients. *Circulation* 2005; 112: 3738–3744.
14. Curtis JP, Sokol SI, Wang Y, et al. The association of left ventricular ejection fraction, mortality, and cause of death in stable outpatients with heart failure. *J Am Coll Cardiol* 2003; 42: 736–742.
15. Cardinale D, Colombo A, Bacchiani G, et al. Early detection of anthracycline cardiotoxicity and improvement with heart failure therapy. *Circulation* 2015; 131: 1981–1988.
16. Chong V, Singh J, Parry H, et al. Management of noncardiac comorbidities in chronic heart failure. *Cardiovasc Ther* 2015; 33: 300–315.
17. https://www.cms.gov/medicare-coverage-database/details/nca-decision-memo.aspx?NCAId=148&NCDId=110&ncdver=3&NcaName=Implantable+Defibrillators&IsPopup=y&bc=AAAAAAAAEAAA& (accessed 6 October 2015).
18. Davis BR, Kostis JB, Simpson LM, et al. Heart failure with preserved and reduced left ventricular ejection fraction in the antihypertensive and lipid-lowering treatment to prevent heart attack trial. *Circulation* 2008; 118: 2259–2267.
19. Maurer M, Adamson P, Costanzo M, et al. Rationale and design of the left atrial pressure monitoring to optimize heart failure therapy study (LAPTOP-HF). *J Card Fail* 2015; 21: 479–488.
20. Rahimi A, York M, Gheewala N, et al. Trends in outpatient transthoracic echocardiography: impact of appropriateness criteria publication. *Am J Med* 2011; 124: 740–746.
21. Garvin JH, DuVall SL, South BR, et al. Automated extraction of ejection fraction for quality measurement using regular expressions in unstructured information management architecture (UIMA) for heart failure. *J Am Med Inform Assoc* 2012; 19: 859–866.

22. Koebnick C, Langer-Gould AM, Gould MK, et al. Sociodemographic characteristics of members of a large, integrated health care system: comparison with US Census Bureau data. *Perm J* 2012; 16: 37–41.
23. Kim SS, Hijazi ZM, Lang RM, et al. The use of intracardiac echocardiography and other intracardiac imaging tools to guide noncoronary cardiac interventions. *J Am Coll Cardiol* 2009; 53: 2117–2128.
24. Nissen SE and Yock P. Intravascular ultrasound: novel pathophysiological insights and current clinical applications. *Circulation* 2001; 103: 604–616.
25. Lang RM, Badano LP, Mor-Avi V, et al. Recommendations for cardiac chamber quantification by echocardiography in adults: an update from the American Society of Echocardiography and the European Association of Cardiovascular Imaging. *J Am Soc Echocardiogr* 2015; 28: 1–39.

Appendix 1

Keywords or phrases used to search the description of ejection fraction and the selected priority in the searching algorithm

1. **First priority**: LVEF, EF, left ventricular (lv) ejection fraction, ejection fraction, left ventricular (lv) systolic function, left ventricular (lv) function.
2. **Second priority**: left ventricular (lv) contractility, systolic ventricular function, systolic function, cardiac function, ventricular function, biventricular function, overall function, ventricular contractility, systolic contractility, overall contractility, ventricular contraction, systolic performance, wall motion, global hypokinesis.
3. **Third priority**: systolic dysfunction, ventricular dysfunction.

Appendix 2

Keywords used to search for ejection fraction conditions

1. **Good condition description**: normal, satisfactory, excellent, good, adequate, intact, preserved, hyperdynamic, hyperkinetic, vigorous.
2. **Worse condition description**: mild(ly), moderate(ly), severe(ly), poor(ly), sluggish, reduce(d), decrease(d), depress(ed), impair(ed), impairment, abnormal, below normal, mild(ly) to moderate(ly), moderate(ly) to severe(ly).
3. **History condition description**: previous, prior, last, recent(ly), history, histories, hx.
4. **Negation condition description**: without, not, no, n't.

Appendix 3. Conversion of text description of ejection fraction into numeric value.

Text description of ejection fraction	Assigned EF value
severe(ly)	16
moderate(ly) to severe(ly)	30
mild(ly) to moderate(ly), mild(ly), poor(ly), sluggish, reduce(d), decrease(d), depress(ed), impair(ed), impairment, abnormal, below normal	45
normal, no abnormal, intact, preserved, adequate, good, satisfactory, excellent	55
hyperdynamic, hyperkinetic, vigorous	70

EF: ejection fraction.