▶ Predicting need for intervention in individuals with congestive heart failure using a home-based telecare system

Elaine Biddiss*[†‡], Simon Brownsell*[§] and Mark S Hawley*[§]

*School of Health and Related Research, University of Sheffield, Sheffield, UK; [†]Bloorview Research Institute, Bloorview Kids Rehab, Toronto, Canada; [‡]Institute of Biomaterials and Biomedical Engineering, University of Toronto, Toronto, Canada; [§]Department of Medical Physics and Clinical Engineering, Barnsley Hospital NHS Foundation Trust, Barnsley, UK

Summary

We have studied how well the need for a medical intervention can be predicted by a telecare monitoring system. During a study period of about 18 months, 45 elderly individuals with congestive heart failure used a home health monitor to enter daily information pertaining to their symptoms and health status. A total of 8576 alerts were generated by the monitoring system, although in most cases, patient and service provider interaction was not required. When system alerts were considered to be serious, or if symptoms persisted, the patient was contacted. A total of 171 key medical events (6 deaths; 28 hospital admissions; 59 changes in medication; 54 cases of advice given; 24 instances where immediate medical attention was recommended) were recorded in the monitoring logs. A multivariate logistic regression model was developed to predict these medical interventions/events. The model correctly predicted key medical events in 75% of cases with a specificity of 74% and an overall cross-validated accuracy of 74% (95% CI, 68–80%). Key predictors included the number of system alerts, self-rated mobility, self-rated health and self-rated anxiety. This suggests that subjective measures are useful in addition to physiological ones for predicting health status. A multivariate decision support model has potential to supplement practitioners and current telecare systems in identifying heart failure patients in need of medical intervention.

Introduction

In the United States, 4.8 million people are affected by chronic heart failure (CHF). The associated health-care costs amount to $38.1 billion annually.[1] Regular monitoring is especially pertinent to the management of CHF where signs of diminishing health may be subtle and difficult to recognize.[1,2] Nevertheless, it has been suggested that many of these symptoms of worsening health (i.e. oedema, dyspnoea, weight gain) present 8 to 12 days before hospitalization.[3] Telecare systems may be a cost-effective and patient-acceptable way of identifying the problems.[3,4]

Despite growing interest and investment in this area, there remain numerous questions as to how to achieve the greatest increase in clinical and cost effectiveness.[5] For instance, questions remain about the best predictors for a particular cardiovascular population.[6] We have therefore studied how well the need for a medical intervention can be predicted by a telecare monitoring system based on self-rated health-related quality of life (HQOL), physical symptoms, lifestyle and physiological measurements in individuals with CHF.

Methods

A review of Barnsley Hospital records identified potential participants for the study. All participants had echocardiographic evidence of heart failure and conventional symptoms. The exclusion criteria were: (1) ejection fraction >40%; (2) unstable angina; (3) age less than 60 years; (4) debilitating dementia or psychiatric disorder; (5) inability to comprehend words presented on an electronic screen; (6) planned coronary revascularization procedures; (7) on a waiting list for heart transplantation; (8) participation in another, conflicting heart failure research study within the past 6 months; (9) lack of an operational home telephone line and a nearby electrical socket; and (10) not living in the mainstream housing sector (e.g. residential or nursing care).

Data collection

Participants were provided with a health monitor (Doc@Home, Docobo Inc, Sheffield, UK) through which they entered daily information pertaining to their symptoms and health status through a set of questions developed by the research and clinical team. Daily measurements of blood pressure and pulse rate were recorded using a wireless monitor (UA-767 Plus BT, A&D Medical). Daily weight was also measured (HCV800, Hanson). Twice weekly, patients also completed a health-related quality of life measure (EQ-5D)[7] directly on the health monitoring unit, giving data on self-rated health (visual analogue scale), mobility, self-care, usual activities, pain/discomfort and anxiety/depression. The data were transmitted at night through telephone lines and screened for abnormalities. If the data fell outside user-specific limits, clinicians were notified. This therefore provided two additional sets of data: (a) a daily record of system alerts generated by the home monitoring system in accordance with the variables listed in Table 1, and (b) a qualitative log of clinical interventions and medical events recorded by the monitoring health-care staff. The study was approved by the appropriate ethics committee.

Data analysis

The data analysis was conducted using standard statistical software (SPSS 14.0). Logistic Regression (LR)[8–12] was used to predict the occurrence of key medical events/interventions as extracted from the health-care practitioner logs. For each week of the study, the average/median values of predictor variables were calculated, the number of system alerts was enumerated, and the presence of a key medical event or intervention was noted from the logs of

Table 1 Default criteria used by the monitoring system to generate system alerts. Specific limits could be set for each individual, as appropriate

Health factor query	Alert generated if...
Daytime shortness of breath	More than usual
Night-time shortness of breath	More than usual
Need for extra pillows	Yes
Swollen ankles	More than usual
Bloated stomach	More than usual
Dizziness	More than usual
Urine excretion	Less than usual + increased dizziness + increased bloating
Cough	New or worse than usual
Weight gain	If weight increases by 2 lb since previous day or if weight increases by 3 lb in a rolling 7 days
Systolic blood pressure	If below 100 mmHg or drops by 20 mmHg from previous reading
Pulse rate	If below 55 per min or above 120 per min
Diet (i.e. eating well)	No or 'less than usual' for a consecutive 7 days
Medication taken	No
Angina	More than usual

monitoring health-care practitioners. The rationale for this approach was two-fold:

(1) Data were collected with varying frequency (either weekly, twice a week or daily);
(2) A time lag was evident between user-inputted data, generated system alerts and nurse responses.

Stepwise, forward selection based on the log likelihood ratio was used to avoid over-fitting the model.[8] The importance of all potential interactions was evaluated via the likelihood ratio test. The model's goodness of fit was assessed based on changes in deviance (i.e. >4 indicating a poor fit).[12]

Models were evaluated using K-fold cross validation with K = 10. Due to the large number of non-events (i.e. a meaningful health-care intervention was not required) in comparison to key events (i.e. a meaningful health-care intervention was required), over-sampling[13,14] was used to obtain a balanced data set. The approximate proportion of key events to non-events was maintained in each test set. The sensitivity (i.e. the proportion of key events that were correctly classified), specificity (i.e. the proportion of non-events that were correctly classified) and overall prediction accuracy were used to evaluate the performance of the model.

Results

Of the 45 participants, six individuals died during the course of the study and eight returned their equipment. The average duration of data collection was 18 months (SD = 5). The patient characteristics are summarised in Table 2.

Predicting the need for medical intervention

A total of 8576 alerts were generated by the monitoring system. In most cases, the response to system alerts did not require patient and service provider interaction. When system alerts were considered of greater severity, or if symptoms persisted, the patient was contacted. A total of 171 key medical events (6 deaths; 28 hospital admissions; 59 changes in medication; 54 cases of advice given; 24 instances where immediate medical attention was recommended) were recorded in the monitoring logs. There were 154 weeks during which one or more key medical events occurred and 2779 weeks during which no key medical events were observed for the participants. Generation of a system alert and subsequent response by a health-care practitioner was not considered a key medical event unless a specific action was actually required and taken (i.e. false alarms were not counted as key medical events). In order to obtain an approximately balanced dataset, key medical events were over-sampled by a rate of 18 times. The average number of medical events experienced per patient per year was 3.5 (with a median of 2 and an interquartile range of 1–4). The average number of non-key alerts generated per patient per year was 49 (with a

Table 2 Patient characteristics (*n* = 45)

Characteristic	Participants (%)
Gender	
Male	83
Female	17
NYHA classification	
2	43
2–3	17
3	40
Age (years)	
60–64	10
65–69	26
70–74	26
75–79	24
80–84	10
85–89	5
Living arrangements	
Alone	26
With partner/spouse	62
More than 2 people in household	12
Smoking	
Non-smoker	82
Smoker	18
Exercise	
None	67
Light	28
Moderate	5
Number of prescribed medications	
<5	13
5–10	59
11–15	23
>15	5

median of 49 and an interquartile range of 47–51). The average percentage of total alerts that were identified as key medical events was 6.4% (with a median of 4 and an interquartile range of 1.4–8).

Table 3 summarizes the univariate significance of predictor variables. All variables with a *P* value <0.1 were

Table 3 Variables considered as predictors of key medical events/interventions based on univariate statistical significance (chi-squared test for binary predictors; independent *t*-test for continuous)

Variable	Scale	*P*
Number of system alerts	Continuous	<0.001
Quality of sleep	Binary (0 = as usual or more; 1 = less than usual or none)	<0.001
Extra pillows needed	Binary (0 = no; 1 = yes)	<0.001
Short of breath (day)	Binary (0 = as usual or less; 1 = more than usual)	<0.001
Diet	Binary (0 = as usual or more; 1 = less than usual)	0.001
Cough	Binary (0 = as usual or none; 1 = new or worse cough)	<0.001
Weight	Binary (0 = within parameters; 1 = outside of parameters)	<0.001
Fatigue	Binary (0 = as usual or less; 1 = more than usual)	<0.001
Self-rated health	Continuous	<0.001
Self-rated mobility	Binary (0 = no problems; 1 = some problems or unable)	0.001
Self-rated anxiety	Binary (0 = none; 1 = moderate or extreme)	0.002
Exercise	Binary (0 = some exercise; 1 = no exercise)	0.002
Self-rated pain	Binary (0 = none; 1 = moderate or extreme)	0.012

examined for inclusion in the logistic regression model in accordance with the guidelines in Hosmer & Lemeshow.[12] Some of the variables were highly correlated (see Table 3). For example, self-rated mobility and self-rated pain were highly correlated (r = 0.74, *P* < 0.01). In these cases, the strongest predictor in the group of correlated variables was selected for inclusion in the model.

The best multivariate logistic regression model for prediction of key medical interventions/events is shown in Table 4. The number of alerts generated by the system emerged as the primary predictor. As shown by the odds ratio (e^{β} = 1.196) listed in Table 4, for every additional system alert generated, the probability of a key medical event increased by 19.6%. On its own, this variable predicted 82% of non-events and 61% of key events with an unadjusted coefficient of β = 0.183. The addition of subjective factors (i.e. self-rated mobility, health and anxiety) improved prediction significantly (log likelihood ratio, *P* < 0.001). The ROC (receiver-operator curve) for this revised model is shown in Figure 1a. With a classification cut-off probability of 0.5, the overall cross-validated prediction accuracy was 74% (95% CI, 68–80). Most importantly, the sensitivity (i.e. prediction of key events) was increased from 61% to 75% with a specificity (prediction of non-events) of 74%. With a maximum sensitivity of approximately 80%, the specificity dropped to 67% (cut-off = 0.62). The sensitivity and specificity of the model for a range of classification cut-off values is shown in Figure 1b. Of the data, 100% of key medical events and 97% of non-events were well-fitted (i.e. change in deviance <4). Approximately 72% of poorly fitted data points were associated with four particular participants. A significantly higher number of daily system alerts (1.5) were associated with these patients as compared to the average (0.5), *P* < 0.001.

Finally, a breakdown of all the daily system alerts (*n* = 8576) generated by the telecare monitoring system indicated that a large proportion of the total number of alerts were attributable to the physiological measurements, namely pulse rate, blood pressure and weight gain. In response to these alerts, monitoring practitioners typically flagged the patient for elevated observation. If the symptom persisted, a telephone call was made to ascertain the health of the patient and possible reasons for the physiological change (e.g. over-eating or an unrelated cold).

Table 4 Logistic regression model for prediction of key medical interventions/events

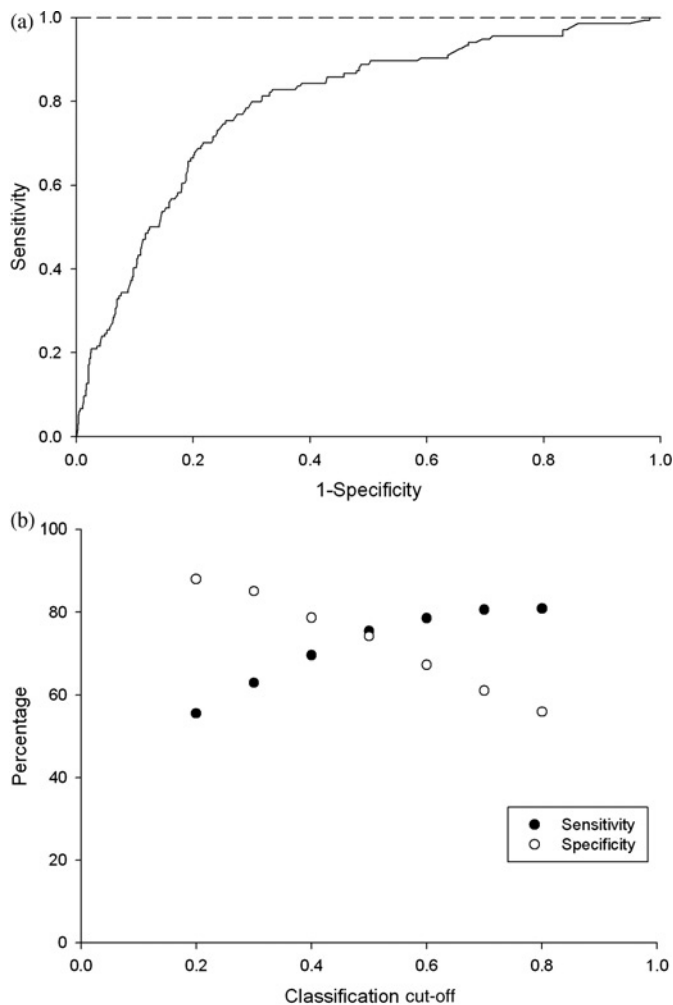| Variables	Coefficient (β)	Standard error	*P*	Odds ratio (e^{β}) [95% CI]
System alerts	0.179	0.008	<0.001	1.196 [1.178–1.214]
Self-rated mobility	0.441	0.093	<0.001	1.559 [1.299–1.871]
Self-rated health	−0.009	0.002	0.012	0.991 [0.986–0.995]
Self-rated anxiety	0.144	0.084	0.19	1.157 [0.982–1.364]
Constant	−0.863	0.216		0.436

Figure 1 Performance of the logistic regression model in predicting key medical interventions/events. (a) Receiver-operator curve (b) Sensitivity and specificity for a range of classification cut-off values

In approximately 86% of cases, alerts generated by these physiological measures were not accompanied by a key medical event/intervention. Alerts pertaining to physical and psychological symptoms, such as anxiety, swollen ankles and the need for extra pillows at night, were most often correct in predicting a key event (Figure 2).

Discussion

The model predicted key medical events/interventions with an overall accuracy of 74% (95% CI, 68–80). With a classification cut-off probability of 0.5, the sensitivity of the model was 75% and the specificity was 74%. To minimize the risk of not identifying the need for medical intervention, a higher classification cut-off probability could be used. This increase in sensitivity is of course accompanied by a decrease in specificity (i.e. more false alarms). The strongest predictor in this model was the cumulative number of system alerts generated in a given week. As can be seen in Figure 2, it is the system alerts stemming from patients' subjective descriptions of their symptoms as opposed to the physiological metrics that are most indicative of the need for medical intervention. This implies that patients are giving medically meaningful reports of their symptoms. With additional predictors based on subjective health perceptions (i.e. self-rated mobility, health and anxiety), correct predictions of key medical events were improved from 61% to 75%. Consequently, in order to increase the effectiveness of telecare systems, it may be important to record more than just physiological variables. Further research to explore a wider range of potential predictors is needed.

The performance of the model's predictions was compared to the clinicians' responses. To estimate the positive predictivity (i.e. the ratio of true key events to the
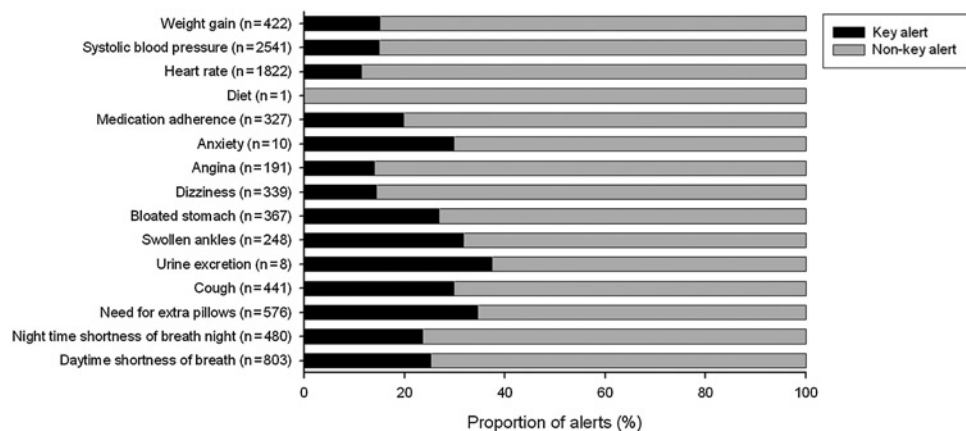

Figure 2 Number of system alerts according to class (e.g. weight gain, increased anxiety, worsened cough). For each class of alert, the proportion associated with a key medical intervention/event is indicated (in black). For example, reports of weight gain generated 422 alerts during the course of the study. In approximately 15% of cases, this symptom was associated with a key medical intervention/event, whereas in 85% of cases, it was not associated with a key medical intervention/event

total number of key events predicted) of the clinicians, the assumption was made that contact with a patient that ensued from a medical concern, but did not result in a key intervention, was a false alarm (i.e. a false positive). Thus the positive predictivity was estimated to be 39% for clinicians as compared to 75% for the decision-support model. Two thirds of all false alarms generated by the model (i.e. incorrectly predicted key events) were instances where the clinician involved also demonstrated a heightened concern for the patient and decided to increase monitoring of and/or contact the patient based on the information collected. This suggests that, although incorrect in its prediction of a key medical event in these cases, the decision-support model did identify instances of elevated risk in line with clinicians' assessments.

Clinical significance

It is important to emphasize that predictive models should be regarded as a tool to assist, not supersede, clinical decision making and prioritizing. Furthermore, 'non-key interventions' (i.e. contact with patients that did not result in a tangible medical intervention), may still have rendered a meaningful health-care service by promoting patient satisfaction, increasing confidence in the quality of care provided, alleviating feelings of social isolation, increasing perceived social support, encouraging adherence to treatment recommendations, addressing a co-morbidity or other problem, and improving clinician-patient relationships. All of these social and health-care perceptions have been implicated as factors in hospital re-admission rates, mortality and/or quality of life in CHF patients.[1,2,15–18] Increasingly, emphasis is being placed on patients with CHF to self-care through initiatives such as the Expert Patient Programme.[19] It may be the case that telecare systems with predictive modelling could complement such initiatives. After all, early identification of high-risk patients, improved home care, and education on heart failure and self-management, are fundamental strategies to decrease morbidity and mortality among patients, and to alleviate the economic burden of frequent hospital re-admissions.[20–22]

Study limitations and future work

The present study explored the development of a decision-support tool that incorporated physiological measurements, physical symptoms and subjective perspectives on health and well-being. We have identified a few important predictors of health status and have explored the factors framing self-rated HQOL. Larger datasets will enable the development of more accurate, robust and generalized models that can predict not only the occurrence of a key medical event and/or need for intervention, but the level of severity of the event.

Logistic regression does not account for the longitudinal nature of the data. It is a 'safe' estimate in that it is more likely to include potentially unimportant variables, rather than excluding important predictors.[23] Although the odds ratios are likely to be similar, standard errors may be underestimated in comparison with methods that account for repeated measures such as generalized estimating equations (GEE).[23] The latter approach, however, requires the assumption that missing data occur completely at random and independently of the outcome variable. In our study, missing data were commonly due to hospitalization and the patient's inability to access their home monitoring system. For comparison purposes, a GEE model was constructed. System alerts and self-rated mobility emerged as the primary predictors with odds ratios similar to those of the simple logistic regression model. As expected, the standard errors calculated through GEE were significantly larger for both system alerts (SE = 0.017) and self-rated mobility (SE = 0.225). The goodness-of-fit measure for the GEE model incorporating system alerts and self-rated mobility was slightly higher than the model which also included self-rated health and self-rated anxiety. This implies that the latter two variables, although not significant predictors, may contribute to the goodness-of-fit of the GEE model.

The necessity of carrying out analyses on a weekly basis to account for time delays between user inputs, system response and clinical action, should also be noted. Inconsistent adjustments of system parameters by monitoring practitioners may also have occurred. For example, for some individuals whose physiological measurements had greater acceptable fluctuations than others, system limits were changed to eliminate superfluous alerts, while for others, these alerts were simply ignored. This may have affected the model fit and issued a higher number of false alarms. Differences in each individual's ability to self-manage (e.g. medication, diet) were not captured in this analysis. It is also possible that some individuals may have been more in tune with their health needs than others and that the model could be refined to reflect patient variations in sensitivity or anxiety regarding perceived symptoms.

Conclusions

Four important conclusions emerged from the study with respect to the performance and development of telecare systems. First, the system for health monitoring used in this study proved useful in indicating when medical interventions are needed. Second, the performance of the systems could be improved by including targeted questions relating to health outcomes. Third, self-perceived symptoms and health status were valuable indicators. Last, a multivariate decision support model has potential to supplement practitioners and current telecare systems in identifying CHF patients in need of medical intervention. Inclusion of such systems in real time could enhance system effectiveness, enable preventative health care and increase practitioner efficiency.

References

1 Zickmund SL, Blasiole JA, Brase V, Arnold RM. Congestive heart failure patients report conflict with their physicians. *J Card Fail* 2006;**12**:546–53

2 Vinson JM, Rich MW, Sperry JC, Shah AS, McNamara T. Early readmission of elderly patients with congestive heart failure. *J Am Geriatr Soc* 1990;**38**:1290–5

3 Clark RA, Inglis SC, McAlister FA, Cleland JG, Stewart S. Telemonitoring or structured telephone support programmes for patients with chronic heart failure: systematic review and meta-analysis. *BMJ* 2007;**334**:942–51

4 Paré G, Jaana M, Sicotte C. Systematic review of home telemonitoring for chronic diseases: the evidence base. *J Am Med Inform Assoc* 2007;**14**:269–77

5 Brownsell S. *Supporting Long Term Conditions and Disease Management through Telecare and Telehealth: Evidence and Challenges*. See http://www.icn.csip.org.uk/nl/?l=202_1_1_3 (last checked 27 March 2009)

6 Artinian N. Telehealth as a tool for enhancing care for patients with cardiovascular disease. *J Cardiovasc Nurs* 2007;**22**:25–31

7 Lenzen MJ, Scholte op Reimer WJ, Pedersen SS, *et al*. The additional value of patient-reported health status in predicting 1-year mortality after invasive coronary procedures: a report from the Euro Heart Survey on Coronary Revascularisation. *Heart* 2007;**93**:339–44

8 Bagley SC, White H, Golomb BA. Logistic regression in the medical literature: standards for use and reporting, with particular attention to one medical domain. *J Clin Epidemiol* 2001;**54**:979–85

9 Wielandt T, Mckenna K, Tooth L, Strong J. Factors that predict the post-discharge use of recommended assistive technology (AT). *Disabil Rehabil Assist Technol* 2006;**1**:29–40

10 Johansson U, Bernspång B. Predicting return to work after brain injury using occupational therapy assessments. *Disabil Rehab* 2001;**23**:474–80

11 Biddiss EA, Chau TT. Multivariate prediction of upper limb prosthesis acceptance or rejection. *Disabil Rehabil Assist Technol* 2008;**3**:181–92

12 Hosmer DW, Lemeshow S. *Applied Logistic Regression*. New York: Wiley, 1989;307

13 Eitrich T, Kless A, Druska C, Meyer W, Grotendorst J. Classification of highly unbalanced CYP450 data of drugs using cost sensitive machine learning techniques. *J Chem Inf Model* 2007;**47**:92–103

14 Kotsiantis S, Kanellopoulos D, Pintelas P. Handling imbalanced datasets: a review. *GESTS Int Trans Comp Sci Eng* 2006;**30**:25–36

15 Yu DS, Lee DT, Woo J, Thompson DR. Correlates of psychological distress in elderly patients with congestive heart failure. *J Psychosom Res* 2004;**57**:573–81

16 Gorkin L, Schron EB, Brooks MM, *et al*. Psychosocial predictors of mortality in the Cardiac Arrhythmia Suppression Trial-1 (CAST-1). *Am J Cardiol* 1993;**71**:263–7

17 Williams RB, Barefoot JC, Califf RM, *et al*. Prognostic importance of social and economic resources among medically treated patients with angiographically documented coronary artery disease. *JAMA* 1992;**267**:520–4

18 Murberg TA, Bru E. Social relationships and mortality in patients with congestive heart failure. *J Psychosom Res* 2001;**51**:521–7

19 Griffiths C, Foster G, Ramsay J, Eldridge S, Taylor S. How effective are expert patient (lay led) education programmes for chronic disease? *BMJ* 2007;**334**:1254–6

20 Lennie TA, Moser DK, Heo S, Chung ML, Zambroski CH. Factors influencing food intake in patients with heart failure: a comparison with healthy elders. *J Cardiovasc Nurs* 2006;**21**:123–9

21 Proctor EK, Morrow-Howell N, Li H, Dore P. Adequacy of home care and hospital readmission for elderly congestive heart failure patients. *Health Soc Work* 2000;**25**:87–96

22 Cline CM, Israelsson BY, Willenheimer RB, Broms K, Erhardt LR. Cost effective management programme for heart failure reduces hospitalisation. *Heart* 1998;**80**:442–6

23 Williamson DS, Bangdiwala SI, Marshall SW, Waller AE. Repeated measures analysis of binary outcomes: applications to injury research. *Accid Anal Prev* 1996: **28**:571–9