Artificial intelligence enabled tele-health coaching changes patients' self-reported patient activation and health outcomes:

A prospective cohort study

Mariana Wieske MSc, Dr Fiona Hamilton, Dr Shoba Poduval, Dr. Joachim Werr

Department of Epidemiology and Public Health

Submitted as partial fulfilment of the requirements for the MSc in Population Health, University College London

August 2019

TABLE OF CONTENTS

LAY SUMMARY

In the last decade, presentations to A&E hospital departments in the UK have increased by 42%. Long-Term Conditions(LTCs) including diabetes, heart diseases, clinical depression constitute 61% of these admissions. Patients who managed their LTCs through continuous treatment, pharmacological or otherwise, had 38% fewer A&E admissions. Management of LTCs may benefit from personalised interventions to prevent admission. The 2019 NHS Long-term Plan mandates better LTC 'self-management' through personalised support and use of technology. In this context, Artificial Intelligence(AI) technologies combined with health coaching interventions can help facilitate identification of patients at greater need of support requiring personalised care and improve self-management.

Not a lot of evidence has been published about the impact of AI-supported interventions on patient health outcomes. From this perspective, a UK cohort study evaluated the impact of a novel AI-supported Pro-active tele-Health coaching intervention.

Health Navigator Pro-active tele-Health Coaching(PHC) utilised AI to identify at-risk-patients to deliver a personalised care programme to reduce hospitalisation. The patients identified through AI 'risk prediction' and selected by clinicians, received support in managing their conditions from a health professional(coach). The intervention consisted of a patient activation programme designed to educate and motivate patients to achieve health improvement goals(patient activation). Patient activation aimed to increase self-confidence and awareness of the condition to minimise emergency hospital admissions.

The PHC intervention was delivered in York in the UK. York adult patients received the intervention between 2015-2019 and the impact was evaluated through collected self-reported activation level and health outcomes data. The York study included 288 patients and was nested within a larger study of 3,000 patients across nine UK areas. Patients were on average 75 years of

age, suffered at least with one LTC and were at high risk of emergency admission.

Two study questions were of interest 1)What was the association between the PHC intervention and general, physical and mental health outcomes? and, 2)Did the AI-enhanced PHC intervention change self-activation in every patient at risk of hospital admission?

The study concluded that, after compared with before receiving the PHC intervention, patients were better able to self-manage their LTCs and experienced improved physical health outcomes.
For self-management, the 33% (from 0.46 to 0.61) per patient improvement in the average activation level score and the 9% change in physical outcomes in the average score from 40.48 to 44.17 (highest=100, lowest=0) were significant. The self-management activation level was also significant in predicting patients' general, physical and mental health outcomes. Sex, age and socioeconomic status indicators (including living environment, housing, employment) were also important factors at some but not all of the levels. For instance, women had a higher activation level and better physical health outcomes compared to men.

As the next steps commissioners may use the study results to extend the evaluation in other UK areas to better understand the relationship between patient activation and health outcomes. Furthermore, future evaluation should include primary care and hospital demographics, diagnosis, service utilisation data and qualitative interviews with health coaches and patients.

ABSTRACT

Background

A&E admissions for UK patients with long-term conditions(LTCs) is increasing. The Artificial Intelligence(AI) enabled Proactive tele-Health Coaching(PHC) intervention reduces admissions. This study hypothesised that patients who were receiving PHC were better able to self-manage their condition and experienced improved general, physical and mental health outcomes.

Methods

Nested within a Randomized Control Trial in UK health settings, this study included the intervention group only and used a before/after design. 855 adult patients received PHC(mean age 75 years, minimum one LTC and at high-risk of emergency readmission) were followed between 2015–2019 in Vale of York(York). The primary interest was the self-reported health outcomes (Short Form12v2(SF12)) and secondary, the impact on patients' ability to manage their health assessed through Patient Activation Measures(PAM13). Quantitative analyses included pre-intervention baseline characteristics, multivariate regression and variances before/after the intervention.

Results

Analysis confirmed a 33% increased PAM13 level (p=0.002, 95% CI0.53-0.67). PAM13 was a significant (p=<0.001) predictor for SF12 General (95% CI0.24-0.41), Physical (95% CI6.39-10.13), and Mental (95% CI5.14-8.48)

health outcomes. Dichotomized and ANOVA tests confirmed a significant change in Physical health (MS=44.16 p=<0.05, 95% CI41.47-46.86), not in Mental (p=0.06) or General (p=0.12) health. Evidence that sex, age, living environment, housing, employment, health disability were impacting factors across levels was insufficient.

Conclusion

The findings partially confirmed the hypothesis. The intervention improved the patient's ability to better manage their conditions and their Physical health only. Limitations in the study were possible causes of these findings. No conclusion could be drawn about the effect compared to usual care.

ACKNOWLEDGEMENTS

I would like to thank the Health Navigator team Eleitha Medina, Geraldine Rook, Moira Docherty Clarke, Andries van der Watt, Amanda Kemp, Elizabeth Robinson, Elizabeth Topps, Nicola Frecknall for working with me and the support in sharing their knowledge, accessing and collating the data.

My UCL supervisors Drs Fiona Hamilton, Shoba Poduval and Health Navigator supervisor Dr Joachim Werr. They critically challenged my thinking and formed my sounding board in resolving obstacles encountered.

Subject matter experts for sharing their knowledge about SF12 and PAM13, research, analytics, Christopher Sherlaw-Johnson of the Nuffield Trust, George Scott of Vale of York CCG, Drs Nina Rogers and Caroline Clarke of UCL.

Dr Thomas Carlisle of Digital Health London and Prof Martin Bobak who at the start of the journey helped me to get me on the way.

Finally, I want to thank my partner Prof Danilo Arnone who has been there all the way, making it possible for me to take a break from my full-time job to focus on the end sprint.

Dissemination policy

Future results from the study may be submitted for publication in peer-reviewed international journals, included in conference presentations, submitted to regulatory authorities and included in internal reports.

1. BACKGROUND

Context

Over the last decade, UK emergency admissions increased by 42% from 4.25 to 6.02 million(1). 61% of patients admitted to A&E had at least one LTC managed through continuous medication or treatment(2). Interventions to help people manage their conditions may reduce or prevent hospital admission. The NHS Long Term Plan(2019) mandates the enhancement of 'supported self-management' particularly of LTCs through better support(3). Patients who managed their health conditions had fewer emergency admissions (38%) and attendances(32%), and were less likely to attend A&E with minor conditions (32%)(1). Artificial Intelligence(AI) technologies could support clinical judgement by facilitating the identification of patients at risk of hospital readmission. Within this context of preventing illness exacerbation and emergency admission, AI technology enables the prediction and identification of patients at high risk of readmission. In addition, intervention such as Health coaching promotes the shift towards more personalised care and reduced hospitalisation.

Digitally enhanced health coaching interventions evidence

The health coaching intervention is an interactive personalised service to educate and motivate patients (also called patient activation) in achieving goals that enhance and improve health(4). Health coaching comprised various formats, targets different populations and typically consists of components

including; 1) Patient identification, 2) Personalised health care planning, 3) Coaching (education, motivation and problem-solving)(5).

1. AI Patient identification and risk prediction

Risk prediction facilitates accuracy in identifying patients (i.e. at high-risk of hospital admission) to target for interventions(6). Machine learning for instance applies risk prediction algorithms, automatic selection techniques and regression methods to stratify patients, and results are typically more precise than traditional tools(8). However, evidence on AI or machine learning effectiveness in facilitating risk prediction related to hospital admission is very limited. One study showed a small effect of risk prediction effectiveness association of OR 1.12 (95% CI1.09-1.15) compared with OR 1.25 (95% CI1.19-1.32) for the risk score using nurse observations(9). A study that reviewed machine learning to predict the prevalence of non-communicable diseases found that prevalence estimates for non-communicable diseases can be reasonably predicted (median correlation 0.88)(10).

2.Personalised health care planning

As an integral part of health coaching the coach and patient agree on goals and actions for managing the health problems(5). This is an effective way to shift from professional-centred decision-making to support the person in taking ownership of managing their condition. Various systematic reviews (19 RCTs, N=10,856) found that this led to small-moderate improvements in some indicators of self-management(PAM13), improved confidence and skills to

manage own health(medication adherence), physical health, mental health(reduced depression SD difference -2.23, 95% CI-2.52 to -1.95)(5)(11).

3.Coaching and tele-Health Coaching

Telephone health coaching studies reported positive results on health behaviours and outcomes. 94%(32/34) of telephone health coaching intervention review studies had positive outcomes supporting the telephone coaching intervention(12). Telehealth coaching is a "*regular series of phone calls between patient and health professional...to provide support and encouragement to the patient, and promote healthy behaviours such as treatment control, healthy diet, physical activity and mobility, rehabilitation, and good mental health"*(13). This intervention has demonstrated significant(p=<0.05) improvements in 15 RCT and 30 review studies, including self-efficacy, health behaviours, physical activity, mental health status, weight management and medication adherence(14)(15).

Impact on self-management and health outcomes

Self-management at a population level reduces the burden of LTCs on the healthcare system and improve population health outcomes(16). Promotion of patient's self-management skills is expected to result in improved levels of activation, avoiding acute exacerbations, better awareness of when to seek medical help, becoming better educated and informed, leading to reduced health care costs (17)(18)(19). Research has furthermore confirmed the associations between self-management and improvement of LTCs. For

example in; pain management, quality of life, health behaviours and status, for patients with the LTCs e.g. depression and type2 diabetes(17)(20)(21).

In summary, a large number of studies confirmed that health coaching had an overall moderate effect in patients with LTCs specifically related to improved self-management and health outcomes. Findings on the effectiveness of self-management on physical and mental health outcomes were mixed or contradictory but generally, this was moderate. However, the lack of published research suggests that the use of AI or machine learning products in health care is novel. They have not yet widely been implemented or evaluated but early findings seem promising. Amalgamating AI technology and tele-health coaching interventions could lead to positive impacts on health outcomes for specific population groups. Health Navigator Ltd provides such a service, called Proactive tele-Health Coaching(PHC).

AI enabled PHC intervention[1]

PHC enables non-clinical telephone support to educate patients on their conditions, plan their care, and navigate complex health and social care systems. It facilitates increased self-management with the aim to prevent disease progression, improve health outcomes and prevent emergency care. Patients receive help with identifying their reasons for contacting health services, and address triggers for urgent access. This includes understanding and managing their condition(s), addressing social isolation and navigating the care system. Over 6-9 months patients work with PHC health coaches

[1] Intervention description used information from Edgren G, etal, 2016, http://www.health-navigator.co.uk/proactive-health-coaching/ and was verified by the lead nurse January 2019

(including nurses, allied health professionals and medical doctors) who provide non-clinical services complementing existing clinical treatment. This includes motivational support, problem-solving, monitoring progress in implementing personalised plans as well as preparing for consultations. Patients receive proactive calls initially daily, reducing to three weekly, weekly, fortnightly, monthly as their risk of unplanned admission reduces. The higher the risk of admission and disease deterioration the higher the frequency of calls.

The intervention is delivered in four phases 1)Patient identification, 2)Intervention delivery, 3)Monitoring and evaluation, 4)Ending future bookings (figure 1).

Phase 1. Patient identification

PHC employed AI predictive risk stratification software installed on the hospital system, identifying patients' at risk of unplanned admission. AI enhanced the impact by identifying the patients who may benefit from PHC more accurately. The AI processed an algorithm which was based on avoidable hospital admission patterns, medical history, patient activities and care events on a daily basis. It predicted patients at risk (1=low,5=high) and low to high(80%-90%) probability to prevent disease deterioration and hospital admission in the next six months. Patients were reviewed by the clinical and the health coaching team via the Health Navigator online system.

Inclusion criteria

Patients were included if they were aged 18 years or older, attended the emergency department in the past six months, considered to be at high risk (score of 5, 80-90% probability) of becoming heavy users of unplanned hospital care in the next six months.

Exclusion criteria

Patients were excluded if they had contact with hospital services within the past twelve months, and the hospital record included: dementia, psychotic disorders, mental disorders caused by substance misuse, terminal cancer, severe hearing loss, language difficulties that required an interpreter, cognitive ability level which was not sufficient for receiving and responding to telephone counselling, no telephone connection, estimated remaining life expectancy <1 year, due or had major surgery in the last six months.

Phase 2. Intervention delivery

PHC aimed to empower patients, created confidence in managing their conditions and find the right care. The intervention was designed to prevent readmissions and delivered the following activities: 1)Coaches got to know the patient during a face-to-face meeting. 2)Reviewed the patient's challenges, experiences, concerns, comprehension of the condition(s) and provided information. 3)Patient-led agreement on goals, values, outcomes, and priorities. 4)Patient-led planning to achieve behavioural change, referral to health services, community or peer support. 5)The coach and patient agreed to regular telephone contact over 6-9 months.

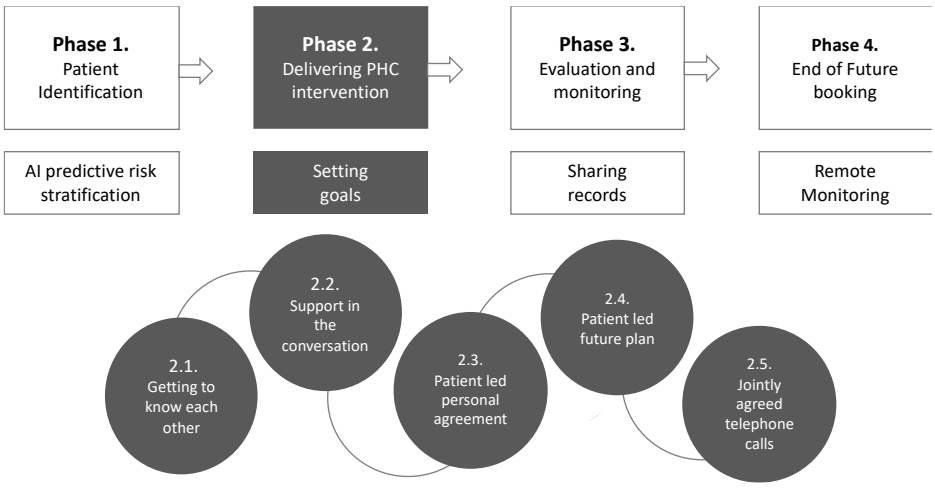
Phase 3. Evaluation and monitoring

Coaches continuously checked progress of the agreed plan. Using the online system, they monitored patient's care actions, hospital admission trends, self-management and health improvement. This was supported by hospital discharge, primary and community care teams through providing specialist services and monitoring.

Phase 4. Ending future bookings

When the patient felt confident and was no longer at risk of admission, a joint decision was taken to not further schedule bookings. This was until the patient either reached out to the coach or experienced deterioration in their health, which may result in hospital admission.

Figure 1. PHC Intervention



PHC Impact

Preliminary findings of the RCT indicated that the PHC intervention reduced hospital admissions of high-risk patients. The RCT furthermore collected self-

management and health outcomes data for the intervention group only. Previous PHC intervention evaluation in a Swedish RCT, 2015 (N=12,181, intervention= 7,280(67%), control=3,290(33%)) showed an overall 12% (CI0.4–0.19) decreased hospitalisation(22). In addition, the intervention was piloted in York, 2016 (N=509, intervention=348, control=161). Results showed a 20% reduction in total health care utilisation(23) and cost, 21% reduction (p=0.01, 95% CI0.94-0.23) in the total number of unplanned admissions and 26% reduction (p=0.01, 95% CI1.23-0.87) in the number of A&E visit[2]. The RTC has since recruited 1,700 patients (out of 3,000 target) in eight sites across the UK.

These study results evaluated the impact of the PHC intervention on hospital utilisation. Understanding what factors contributed to these outcomes helped in the scale-up of the intervention. This required a deeper comprehension of the PHC intervention impact on patient's PAM13 Level and SF12 general, physical and mental health.

PHC analysis framework

The analysis framework for this study was developed based on the available research evidence and the PHC preliminary evaluation findings. It applied the following logic: The AI technology enabled PHC by accurately targeting patients most at risk based on multiple indicators. Patients were offered coaches who trained them to manage, solved problems and pursued goals. This facilitated behaviour change (i.e. exercise, medication adherence),

[2] Hospital data from York hospital Trust, Health Navigator analysis 2018

increased psychological functioning (depression), and disease control(20). It supported patients actions to maintain self-management as a strategy in meeting objectives(16). Better self-management skills were expected to result in a patient who was more engaged in their health care, better able to manage their disease, avoiding acute exacerbations(18). As the patient became more aware when to seek medical help they learned how to use health services, leading to reduced hospital utilisation(19).

The framework consisted of five components: 1)AI risk prediction technology identified the right patient who would benefit from health coaching 2)Health coaching educated and trained the patients 3)Patients improved self-management 4)Impacted general, mental and physical health disease management, which 5)Reduced hospital readmission and utilisation (figure 2).

The evaluation included components 2 to 4 to develop a deeper understanding of the impact of the PHC intervention on the patient's self-management and their general, physical and mental health. On the basis of existing research, it was expected that the intervention would positively impact the high-risk patient groups. The results further helped to understand the transferability of the intervention in different contexts and UK populations. More generally, this study added to developing population health knowledge of the impact of AI-enhanced interventions on health improvements.

Figure 2. Model for analysis

2. AIMS, OBJECTIVE AND HYPOTHESIS

Aims

This study determined how the AI-enabled PHC intervention impacted a population's health by analysing self-reported SF12 health outcomes and PAM13 patient activation data collated between 2015–2019 in York. It used the PHC RCT data of the intervention group to create an insight into how self-management was linked with the improved health outcomes.

Objectives

The contribution to existing knowledge base led to increased understanding of the intervention outcomes in the UK population with the objectives to 1) Determine if the intervention ultimately resulted in improved SF12 general, physical and mental health outcomes through increased PAM13 patient activation and 2) Identify the association between general, physical, mental health, self-management outcomes and socioeconomic characteristics.

Hypothesis

Previous studies have suggested that health coaching was associated with improved self-management and improved physical and mental health outcomes which led to reduced hospital admissions and costs(16)(18)(19). This study hypothesised that patients received the intervention were better able to self-manage their condition and experienced improved general, physical and mental health outcomes.

Research question

The study evaluated the association between the AI-enhanced PHC intervention and health outcomes for adult patients identified to be at high risk of hospitalisation before/after the intervention at six or eighteen months. In relation to this, the questions the study sought to answer were: 1) What was the association between PHC intervention and general, physical and mental health outcomes measured through the SF12 questionnaires? and, 2) Did the AI-enhanced PHC intervention improve self-activation in patients at risk of hospital admission measured through the change in PAM13 activation level?

3. METHODS

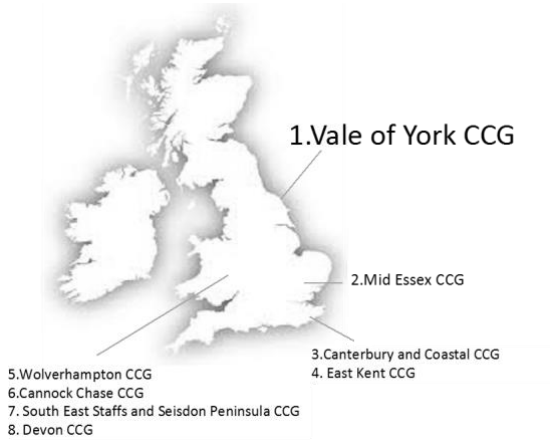
Study design

The cohort study was nested in the multi-centre parallel, two group RCT replicated from Sweden in UK health settings in 2015. Patients were randomised to receive either the intervention or usual care in a 2:1 ratio applying the Zelen's study design. The PHC RCT study only collated self-reported health outcomes and self-activation scores for the intervention group. This decision was based on the high cost and lack of budget to collect data. Therefore, it was anticipated that no comparison could be made with the control group and no conclusion would be drawn about effect compared to usual care. For this reason, the research design was a longitudinal prospective nested cohort study using data from the intervention group SF12 and PAM13 outcomes. Data were collected at baseline before and after the intervention (at six, and/or eighteen months).

Setting

The RCT sample included patients that met the inclusion criteria in eight sites across the UK however, this cohort study analysed data for York only. Patients recruited between 2015-2019 in the intervention and control groups, were followed for up to two years for the occurrence of the hospital utilisation outcomes, starting from the date of randomisation. Following the York analysis, methods will be iterated and refined with the purpose to replicate the study in the remaining sites (figure 3).

Figure 3. Setting



Patient recruitment

The selected patients included high-risk patients that were automatically identified through the risk-stratification AI software tool. Study eligibility was based on the inclusion and exclusion criteria as described in the earlier section; patients were then manually screened by the clinical team and allocated a coach who contacted them for a baseline face-to-face meeting. Upon consent, the SF12 and PAM13 questionnaires were completed during the baseline face-face-meeting. This was repeated at six and/or eighteen months, the patients received the self-reported questionnaires in the post, completed and returned it to the research team.

Outcomes and variables

Data were collated on the population's experienced of health using the SF12 and PAM13 survey tools. Both tools have been found to be a valid, reliable and used for patients with LTCs (17)(24)(25)(26)(27)(28). The primary study outcome was; whether the intervention improved health, as indicated by 1)SF12 general, physical and mental health outcomes measures and

secondly confirmed a change in the ability to self-manage as indicated by the 2)PAM13 tool and 3)Risk factors and predictors. (appendix 1, table 1)

1.SF12 Health outcomes

The analysis of the SF12 Health outcomes identified what the association between PHC intervention and general, physical and mental health outcomes were. The self-reported SF12 measures consisted of twelve descriptive items divided into eight domains: General health, Physical Functioning, Role Physical, Role Emotional, Mental Health, Body Pain, Vitality and Social Functioning. Patients scored their health on the day for each domain. The scores ranged from 1 to 5, 1 indicated that 'they had no problems' in that domain and 5 indicated that 'they had extreme problems'(17). Improvement in health followed the categories of self-perceived health before/after the intervention and was estimated through the 'General health' item of the questionnaire. General Health binary categories were defined as 'good' by answers 1) Excellent, 2) Very Good or 3) Good, and 'bad' by answers 4) Fair or 5) Poor. Physical Health was measured by composite scores including General Health(GH), Physical Functioning(PF), Role Physical(RP) and Body Pain(BP). Mental Health was measured by composite scores including the Vitality(VT), Social Functioning(SF), Role Emotional(RE), and Mental Health(MH).

2.PAM13 outcome

The PAM13 analysis confirmed whether the AI-enhanced PHC intervention improved self-management in patients at risk of hospital admission. This was

measured through the PAM13 questionnaire and the change in the activation level (1=low to 4=high). PAM13 consisted of 13 patient statements related to their health care, managing health related tasks, and self-assessed knowledge of conditions, treatment options, self-efficacy, activation, confidence, competence, and the ability to access relevant support (5)(17). PAM13 measured the changes in these areas as described by Blakemore et al.(2016)(17). Patients agreed or disagreed with the PAM13 statement on a response scale of 1-5 (1=strongly disagree, 4=strongly agree, 5=not applicable). Responses generated a continuous variable from 0-100 the higher scores the more the patient was activated. “5=Not applicable” answers were treated as missing. The total score was generated where patients had answered minimal 10/13 questions. The scores were categorised into four levels for descriptive analysis: Level 1=‘passive recipients of care who do not understand that they can play an active role in their own healthcare’; Level 2=those ‘who lacked the basic knowledge and confidence to effectively self-manage’; Level 3=those ‘with basic knowledge about their health, but they lack the confidence and skills to engage in positive self-management behaviours’; Level 4=patients with ‘the knowledge and confidence to self-manage but who may need support during times of personal stress or health crisis’.

3.Risk factors and predictors

The study was furthermore interested in identifying how SF12 and PAM13 outcomes were associated with population health risk factors and predictors**.** Previous research has identified sex, age, education and diagnosed

conditions (i.e. asthma, COPD, IBS, heart failure) to be risk factors(5)(11)(17)(20). Therefore, available data on sex and age, socioeconomic status UK Index of Multiple Deprivation(IMD) 2015 decile scores, were analysed. IMD is a set of measures identifying aspects of deprivation not affluence. The variables included income, employment, education, health disability, crime, housing and living environment IMD decile scores 1=low to 10=high.

Statistical methods and analysis

Sample size

Across the eight sites the target of 3,000 (to be recruited by 2020) high-risk patients was randomised at 2:1 ratio to receive intervention (n=1,800) or regular care controls (n=900). Power of 90% and alpha α=0.048 enabled a standardised effect size of 0.25 on continuous outcome measures. For York, this approach resulted in a sample size of N=855 (intervention n=591, controls n=263). To maximise measurement power for measuring the outcomes, PAM13 literature recommended a minimal sample size of 200 for the intervention group(25).

Data integrity and quality

For the analysis, the mean with standard deviation, percentage of missing data, and SF12 and PAM13 percentage of ‘not applicable’ answers were considered. Internal consistency was measured as the Spearman correlation coefficient between summary scores, individual scales and different variables.

Similar studies confirmed that strong correlations are rare and correlation r≥0.50 is considered strong, r ≥0.30 moderate and r ≥0.10 weak(26). In the sample the distribution of self-reported health outcomes adjusted mean and overall scores was described by patient activation(PAM13) level, gender and age before/after the intervention. For cases that were excluded due to missing data, differences between included and excluded cases at baseline were tested for similarity. Based on the results an appropriate strategy was applied for accounting for missing data.

Descriptive statistics

Descriptive statistics described the distribution of SF12 and PAM13 outcomes, and changes in scores before/after the intervention. The analysis was based on intention-to-treat completed cases and a specifically developed analysis plan.

As the primary outcomes of interest for this study was the self-reported SF12 Health outcomes, complete cases were defined as patients who completed the questionnaire at baseline and post-intervention, without missing data for the SF12 primary outcomes, with available completed baseline data sex and age. SF12 and PAM13 outcomes were treated as continuous, categorical and normally distributed. Baseline values of outcomes and pre-specified covariates included: SF12 Health outcomes (comprised General Health, Physical Composite Scale, Mental Composite Scale mean scores) and PAM13 self-management mean score(28)(27). The mean ± standard deviation for continuous variables or n(%) for categorical variables included gender and age (categorised 20-29 to 90-99) and IMD socioeconomic decile scores(29).

Regression analysis

The associations between baseline SF12 Health outcomes and PAM13 item scores and independent variables were investigated applying multiple linear regression to compare the change in outcomes before/after the intervention. The correlations coefficients between the variables and risk factors, and impact of potential predictors and risk factors were explored. Bivariate relations were examined prior to the regression analysis applying the Mantel-Haenszel test for homogeneity by matching interactions between the intervention group SF12, PAM13 variable and the variables for sex, age, socioeconomic and time period before/after the intervention. Continuous and categorical data were used in the multiple regression analyses model with independent variables regressors as predictors of change in the outcomes scores after the intervention between the baseline and follow-up. To analyse the mean differences from baseline before/after intervention, paired *t*-test for the total scale was performed and Wilcoxon signed-rank test for individual items.

Hypothesis testing

The hypothesis was tested applying the Wilcoxon test for the null hypotheses in relation to changes before/after the intervention outcomes. Including individual items and subscales to examine individual-level predictors e.g. IMD socioeconomic status, sex and age. Logistic regression calculated the odds of having poor self-management, mental, physical functioning as reported by patients. To compare the changes in the intervention group, analysis of

variance (ANOVA) was performed. Effect sizes used Cohen's interpretation: small(≥0.20), medium(≥0.50) and large(≥0.80)(30).

Analysis of outcomes

The analysis of the outcomes aimed to 1) determine if the intervention improved SF12 Health outcomes and PAM13 self-management outcomes and 2) Identify the association between SF12 general, physical and mental health and PAM13 self-management outcomes and population health risk predictors. Stata version 15 was used in the analysis.

The formal licences for the SF12 and PAM13 questionnaire scoring methods were not purchased. For this reason the researcher developed scoring approaches using existing publicly available protocols, research findings and methods(24)(27)(28)(31)(32)(33)(34). Based on this an Excel spreadsheet was created to collate, score and calibrate the individual SF12 and PAM13 items.

Addressing bias

Selection bias was managed by utilising the data collected building upon the initial RCT study design approach. This approach addressed sequence generation, allocation concealment, baseline comparability, intention-to-treat and loss-to-follow (35).

Selective outcome reporting was managed by declaring upfront that the research was conducted on behalf of ‘for-profit’ companies Health Navigator

AB and Health Navigator Ltd. The Nuffield Trust acted as the Chief Investigator(CSJ)[3], was blinded to the outcomes and received funding of £30k from Health Navigator. The researcher (MW)[4] conducted the analysis and was external to the study. It ensured the research was conducted in line with the approved research protocol.

Recall bias, was not applicable as patients were asked at the point in time about their status of measurement rather than information referred to past events.

Data management

The patient self-reported health outcomes measures were gathered and recorded from questionnaires at baseline, six and/or eighteen months in a secured online database. Patient demographic baseline data were collected via the questionnaires and hospital database system. The socioeconomic status was generated by mapping the IMD scores against the patients' recorded postcodes. All data requested via the research team were made available and accessible to the researcher, who signed a confidentiality agreement. The data were transferred to the researcher according to the confidentiality agreement and NHS ethics regulations, for further data preparation, coding, scoring and calibration. The researcher oversaw the data input progress, providing weekly reports on interim results to maximise a minimum level of data quality. In the case of inaccuracies, missing data, or errors the data were returned to the research team for correction where

[3] Chris Sherlaw-Johnson
[4] Mariana Wieske

possible. The researcher collated pseudonymised hospital data from the hospital databases and matched this against the SF12 and PAM13 questionnaires data. In addition, the researcher conducted four interviews with health coaches to validate the intervention. Due to restricted timelines, hospital and interview data were not incorporated in this thesis.

Ethics

Ethical approval has been obtained from the NHS Research Ethics Committee IRAS Project ID 173319 and NIHR Number CPMS 19857.

4. RESULTS

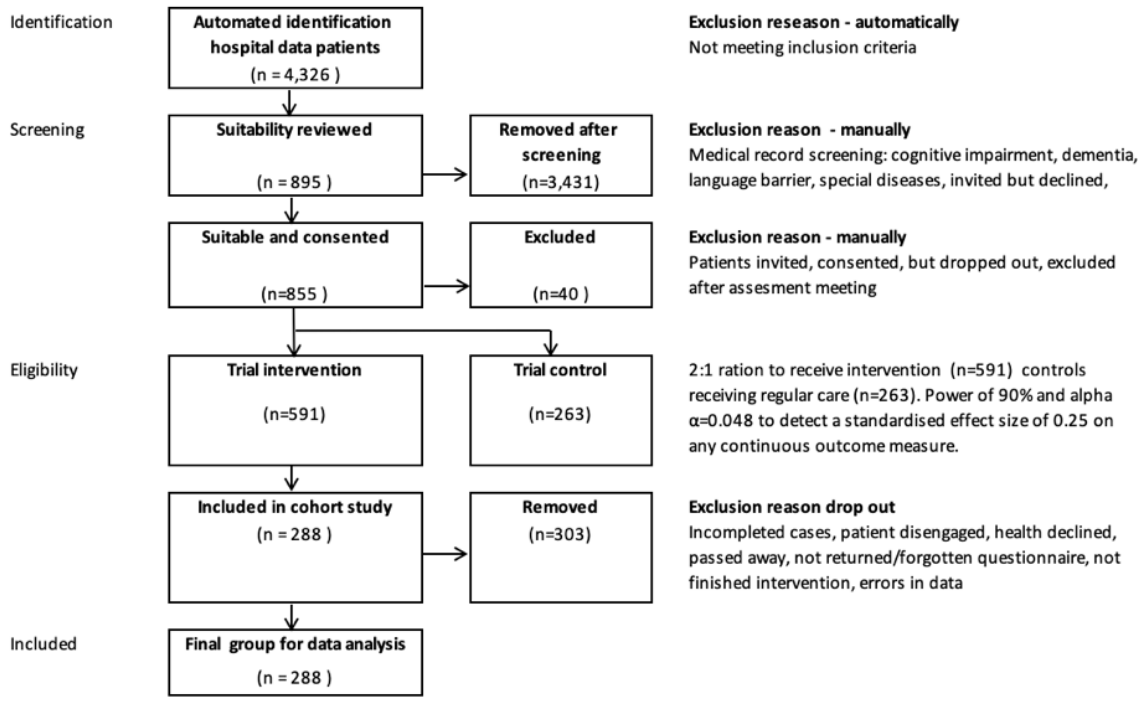
Sample and data quality assurance

The eligible consented sample included 855 patients. 591 subjects underwent the intervention and 264 were controls (figure 3). For SF12 and PAM13, 303/591 (51%) questionnaires were classified as ‘incomplete’ and were excluded. Reasons for exclusion included data recording errors, missing data (≤3 missing items), duplication, non-returned or incomplete questionnaires. Causes for missing data included patients disengaging from services due to declining health/death, withdraw of consent to contact, and moving out of the area. For SF12, 288/591 (49%) of the recorded SF12 questionnaires were eligible for inclusion. For PAM13, 318/591 (50%) were eligible and after calibration against the 288 complete cases, 271 (78%) cases were included. Of the 288 complete cases, SF12 reported 0% item responses missing, and PAM13 between 27%-28% of all 13 individual item responses. There were no missing data for the baseline data sex and age and only 2% across socioeconomic status variables.

For this study ‘complete cases’ analysis with 49% of the full eligible sample, was the most appropriate approach as per the regression and T-test results.The tests confirmed the baseline difference for the included complete(n=288) versus excluded incomplete(n=303) cases for variables Sex (p=<0.001,95% CI -0.85-0.63), Education (p=0.02,95% CI -0.94-0.05) and Crime (p=0.05,95% CI -0.87-0.002). Both groups seemed evenly distributed for Sex, 50% were women in the included group and 57% in the excluded

group. The trend showed a slightly higher proportion of high-level IMD deprivation in Education and Crime in the excluded group (27%) compared with the included group (17%-20%). In the excluded group 47% of Sex data and 17% of Education and Crime data was missing (appendix 2.2). Available data were insufficient to confirm the 'missingness' and the impact on the validity of the final results for the full cohort. The reason for the missing data may be due to input errors rather than related to the intervention. Therefore, it was anticipated that strategies including imputation may not result in significant differences to the study outcomes.

Figure 3.Patients for inclusion



Descriptive statistics

For the 288 participants, the mean age was 75 years (range: 20-99 years), 50% were females. For IMD deprivation items the majority (78%-84%) were at

the ‘least deprived’ end of the scale, for Housing and Environment the least deprived scores were slightly lower (60%-63%). On a binary scale, SF12 General Health was ‘Poor’ for 55% and ‘Good’ for 45%, the PAM13 Level was ‘Low’ for 39.5% and ‘High’ for 60.5% of the patients (table 2).

Distribution was described by SF12 and PAM13 levels, and variables time period, sex, age and IMD overall scores at the baseline level. In all variables for the SF12 and PAM13 outcomes, skewness was <=1.0 and kurtosis <=2.0 this was reasonable. Age was normally distributed with kurtosis=4.4 which showed a slightly heavier tail or outliers (appendix 2.figure1).

The majority of patients scored their SF12 General health to be ‘Fair’ 43% or ‘Good’ 34% and there was a strong (p=<0.001) association with the PAM13 levels (χ^2=69), and Age (χ^2=63) (table 3). Across the four PAM13 levels, data were equally distributed (23%=level 1, 24%=level 2, 30%=level 3, 22%=level 4). Across the PAM13 levels, there was a particularly strong (p=<0.001) association for all SF12 Health outcomes including General (χ^2=39.8), Mental (χ^2=102) and Physical (χ^2=36) health. Testing for baseline differences over a time period, SF12 Health outcomes confirmed a significant (p=<0.005) association with Mental health (χ^2=80) and PAM13 Level (χ^2=12)(table 4).

Associations between PHC intervention and SF12 Health outcomes

The association between PHC and general, physical and mental health outcomes were measured through the SF12 questionnaires. The correlations

between SF12 Health outcomes and PAM13 Levels showed that the associations were of varying strength, but in general, these were moderate(r=0.39-0.34). The strongest(p=<0.001) associations were observed between PAM13 and SF12 Physical (r=0.39), Mental (r=0.36) and General(r=0.34) health. Associations between PAM13 Level and Employment, Health disability, and Environment, between Physical health and Sex, Environment and between Mental health, Sex and age were significant (p=<0.001) but weak (r=-0.10-0.16) (appendix 2.4).

Risk factors, Odd Ratios(OR) and Confidence Intervals(CI) for the associations between SF12 Health outcomes and PAM13 levels and risk factors after the intervention showed likely improvements after the intervention. Generally, patients with higher PAM13 levels were significantly (p=<0.001) more likely to live in good General, Mental and Physical health. They further experienced less deprivation related to health disability and also the environment they lived in. More specifically, after the intervention older patients who lived in less deprived environments had better General health outcomes (p=0.03,95% CI1.01-1.36). They also had more chance of having a higher PAM13 level (p=0.002, 95% CI2.02-4.68).

The study took into account the potential confounding effects that impacted PAM13 Level and Health outcomes. The most significant (p=≤0.03) were identified for General (OR=3.0, 95% CI1.27-7.15) and Physical (OR=1.03, 95% CI1.00-1.06) health. No potential confounders or effect modifiers were identified for Mental health outcomes. For SF12 General health time period was identified as a confounder.

Results showed that controlling for time period, patients were 2.99 times (p=0.002, 95% CI0.37-0.81) more likely to have better outcomes. ‘Physical health outcomes’ potential confounders included time period, age, sex, health disability, housing, environment, and employment. After controlling for these variables, the PAM13 Level OR remained unchanged. Therefore, these variables only played a small role in reducing the effect between PAM13 Level on Physical Health (appendix 2.5,2.6).

PHC impact on PAM13 Patient activation

The analysis results confirmed the AI-enhanced PHC intervention improved activation(PAM13) in patients at risk of readmission. The multivariate regression and ANOVA before/after the intervention results confirmed an association for PAM13 level and Physical health and improved scores after the intervention. The General and Mental health association between ‘poor’ and ‘good’ health, after the intervention, confirmed no statistically significant difference in mean scores. All results were significant (p=<0.05) and trustworthy as indicated by non-violation of the equality of variances Barletts's test (p=0.27-0.53)(table 5).

Comparing the PAM13 levels before/after the intervention, the two-sample test with equal variances showed a positive association. It confirmed an improved score after the intervention (from 0.45 to 0.60, 95% CI0.53-0.67) and increase in individual PAM13 levels mean score (from 2.36 to 2.69, 95%

CI2.55-2.84). The overall mean PAM13 level (23.16) was higher after compared with before (22.40) the intervention, although significant the overall effect size in the difference was very small. The association between Physical health and the intervention confirmed a statistically significant increase after(MS=44.16, 95% CI41.47-46.86) compared with before(MS=40.48, 95% CI37.87-43.08) the intervention but the correlation was weak (r=-0.08).

PAM13 as a predictor for health outcomes

The study hypothesised that patients who receive PHC were better able to self-manage and experience improved general, physical and mental health outcomes measured through the PAM13 and SF12 questionnaires. The analysis evaluated a change in activation level, applying multiple regression including predictors of SF12 Health outcomes at baseline.

Earlier tests indicated significant associations between the SF12 Health outcomes and PAM13 level. Therefore, the regression estimated the numerical relationship using PAM13. The results confirmed that PAM13 Level predicts the SF12 Health outcomes. The linear regression level showed significant (p=<0.001) relationships for time period before/after the intervention (95% CI-0.12 to -0.03), General (95% CI0.24-0.41), Physical (95% CI6.39-10.13), and Mental (95% CI5.14-8.48) health (table 6).

Risk factors predicting SF12 Health outcomes and PAM13 Level were tested through multiple regression analysis. The analysis was interested in

identifying whether the change in PAM13 Level scores associated with SF12 Health outcomes could be due to time period, sex, age, environment, health disability, or employment. By fitting these variables as predictors in the linear multiple regression model, it confirmed that only sex and age were significant impacting factors resulting in three models (table 7).

Model 1.SF12 General health, PAM13 and age

Results confirmed that for each increase in PAM13 Level patients had a better General health outcome compared to those with a lower PAM13 level (coef=0.33, p<0.001) after accounting for age. After adjusting for PAM13 Level and age and it was found that per each year increase in age, there was an increase in General Health (coef=0.18, p=<0.001). The total variability explained by these variables were low (R2=15.78%).

Model 2.SF12 Mental health, PAM13 and sex and age

Patients with a higher PAM13 level had a better Mental health outcome compared to those with a lower PAM13 level (coeff=6.36, p<0.001) after accounting for sex and age. After adjusting for PAM13 Level, sex and age there was an improvement in Mental Health scores (coeff=0.73, p=<0.001). For each level increase in PAM13 Level, Mental health increased by 1.50 (p<0.001), after controlling for age and sex. Those with higher PAM13 level had higher scores compared to those in lower PAM13 levels, both coefficients were significant, holding age and sex, constant.

Model 3.SF12 Physical health, PAM13 and sex

People with a higher PAM13 level had better Physical health outcomes compared to those with a lower level score 6.36 (p=<0.001) after accounting for sex. After adjusting for PAM13 level and sex there was a significant(p=<0.001) 8.11 increase in Physical health score. For each level increase in PAM13 level, the Physical health increased by 1.50 (p=<0.001), after controlling for sex. In particular, women had higher Physical health scores compared to men by 11.53 (p=<0.001,95% CI19.94 -36.50) after accounting for PAM13 level. The total variability in Physical health explained by sex was low(R2=15.12%).

The variables did not show potential effect modification on SF12 Health outcomes and PAM13 Level, they were therefore regarded as potential confounders. The results were adequate, but inspection of the residuals versus the fitted values suggested that there might be high-level heteroskedasticity. The distribution of the errors had some kurtosis and skewness, this was accepted as being reasonable (appendix 2.figure 2).

5. DISCUSSION AND CONCLUSION

Discussion

This study has created insight into the impact of AI-enabled tele-Health Coaching on a population's health. It hypothesised that patients who received the intervention were better able to self-manage their condition and experienced improved general, physical and mental health outcomes. The findings confirmed that the PHC intervention improved PAM13 activation in patients at risk of hospital admission and SF12 physical health.

The cohort were older, high-risk patients, with LTCs. Based on these characteristics, it was expected that patients would deteriorate over time, therefore minimal health related improvement was anticipated. The study explored SF12 outcomes associated with PAM13 patient activation scores and socioeconomic demographics. Patients showed significant improvement in the PAM13 level which included knowledge, skills, and confidence for self-management. The most prominent factor associated with Health outcomes was the PAM13 level in the intervention group of patients of the RCT. Evidence showed that increased activation was associated with positive change in SF12 outcomes.

Associations between PHC intervention and SF12 health outcomes

The association between PHC intervention and general, physical and mental health outcomes were measured through the SF12 questionnaires. The correlation, OR and CI tests confirmed significant but weak correlations between PAM13 Level and SF12 General, Physical and Mental health

outcomes. Overall, patients with higher PAM13 levels had a higher chance of better SF12 Health outcomes. The Mantel-Haenzel test verified that the confounders between PAM13 Level on SF12 Physical Health played a small role in the effect on the outcomes. It was confirmed that the associations between the PHC intervention and all SF12 health outcomes were determined by PAM13 levels. In other words, patients who received PHC were better able to self-manage their conditions and experienced improved SF12 Health outcomes.

Impact of PHC on self-activation

The study findings confirmed that the PHC intervention improved self-activation in patients at high-risk of hospital admission. This was measured through the PAM13 questionnaire and the change in activation level. The multiple regression analysis confirmed that patients with a higher PAM13 level experienced better health outcomes when controlled for sex and age. In particular female patients with higher PAM13 Level compared to men had better Physical health outcomes, both coefficients were significant, keeping sex and PAM13 level constant.

Impact of PHC intervention on self-activation and health outcomes

In addition to the previous dichotomized tests for associations between the intervention and PAM13 and SF12 outcomes the ANOVA test again confirmed that the intervention improved patient's ability to better manage their conditions and their Physical health (MS=44.16 p=<0.05, 95% CI41.47-46.86). The results show a weak correlation for all indicators and do not

support improvements in General health or Mental health. Therefore, the conclusion was that the study findings only partially confirmed the hypothesis. These results will need to be interpreted with caution and no conclusion could be drawn about the effect compared to usual care.

Study strengths

There were only limited published data on IA-enabled interventions and no research specifically related to pro-active personalised health coaching and the impact on health outcomes. This longitudinal study included a large sample size, validated patient-reported measurements allowing assessment of change in patient activation levels and health outcomes.

Limitations

The limitations of the study created insight in the transferability of the intervention in different contexts and populations across the UK.

Replicability - The PHC RCT study outcomes for self-management and health outcomes scores were not collected for the control group. Therefore, a before and after was most suitable although less robust. No comparison could be made with the intervention group and no conclusion drawn about the effect compared to usual care. Secondly, formal SF12 and PAM13 licences were not purchased and scoring mechanism were designed using available published methods and approaches. This may have consequences on the results and the accuracy could not be guaranteed, further restricting generalisability and replicability. Thirdly, the cohort were older patients identified in a hospital setting for one area in the UK. The results may not

generalise the results and replication would be done with caution. Lastly, the study did not confirm the impact of patients' clinical diagnoses which were identified as an important predictor in the literature.

Selection bias - the cohort achieved a 49% response; this may lead to non-response bias. The data were therefore not a strong basis for representing the population, and as such the proportions at each level of activation should be used with caution. As this study only included 'complete SF12 cases' the completion of follow up on SF12 health outcomes measures was 100% and thus potential bias here was less. The PAM13 self-activation completion was 75%, and some scales did suffer additional missing data.

Statistical power - The 'complete cases' analysis, reduced the sample size which may have consequently lost the precision of the estimates impacting statistical power.

Study implications

Within the context of preventing illness exacerbation and A&E readmission avoidance, PHC utilised AI technology to enable personalised care resulting in improved health outcomes in certain population groups. The assessment of socioeconomic factors, validating the impact of the PHC intervention, could inform decision makers and commissioners in the scale-up and focus on targeting wider population groups beyond the hospital population.

Conclusion

This study concluded that 1) people who received PHC were better able to self-manage their LTC and experienced improved physical outcomes. 2) PHC intervention improved PAM13 self-activation in patients at risk of hospital admission. 3) The associations between PHC intervention and General, Physical and Mental health outcomes before/after the intervention was significant.

Future research implications

Due to restricted study timelines, additional received hospital patient data including diagnosis and hospital utilisation were not included in the final analysis. Assessing the relationship between this data and patient activation and health outcomes will be required. Future research should include qualitative interviews with health coaches and patients to validate their views on the intervention impacts. This may strengthen the evaluation of the effectiveness of AI predictive medicine in association with patient health outcomes. Lastly, this cohort study may form the basis for extending the study to the remaining seven geographical areas.

Table1.Outcomes measures

Outcomes	Measures	Instrument	Assessment time points
Baseline	Age, gender, IMD	Self-reported via SF12 and PAM13 IMD measures	Before intervention
Primary: *Health*	Physical Composite Scale	SF12 items:1-General Health, 2a,b-Physical Functioning, 3a,3b-Role Physical, and 5-Body Pain.	Before/after intervention
	Mental Composite Scale	SF-12 items: 6b-Vitality, 7-Social Functioning, 4a,b Role Emotional, and 6a,c Mental Health	
Secondary: *Self-management*	Self-activation	PAM13 items: include depression, health literacy, social support	

Table2.Baseline before/after intervention

Exposure variable	After (n=288)	Before (n=286)	Total (n=574)	Before %	After %
Sex	**288**	**286**	**574**	**50%**	**50**
Female	145	145	290	50%	51%
Male	143	141	284	50%	49%
Age	**288**	**286**	**574**	**50%**	**50%**
20-29	1	1	2	0.3%	0.3%
30-39	2	2	4	0.7%	0.7%
40-49	3	3	6	1%	1%
50-59	16	16	32	6%	6%
60-69	56	55	111	19%	19%
70-79	108	108	216	36%	38%
80-89	89	88	177	31%	31%
90-99	13	13	26	5%	5%
SF General health	**288**	**286**	**574**	**50%**	**50%**
Poor	158	175	333	55%	61%
Good	130	111	241	45%	39%
PAM Level	**195**	**226**	**421**	**34%**	**34%**
Low	77	123	200	40%	54%
High	118	103	221	61%	46%
Deprivation	**283**	**281**	**564**	**49%**	**49%**
Most deprived	45	45	90	16%	16%
Least deprived	238	236	474	84%	89%
Income	**283**	**281**	**564**	**49%**	**49%**
Most deprived	45	45	90	16%	16%
Least deprived	238	236	474	84%	84%
Employment	**283**	**281**	**564**	**49%**	**49%**
Most deprived	44	43	87	16%	15%
Least deprived	239	238	477	84%	85%
Education	**283**	**281**	**564**	**49%**	**49%**
Most deprived	62	62	124	22%	22%
Least deprived	221	219	440	78%	78%
Housing	**283**	**281**	**564**	**49%**	**49%**
Most deprived	113	112	225	40%	40%
Least deprived	170	169	339	60%	60%
Environment	**283**	**281**	**564**	**49%**	**49%**
Most deprived	107	105	212	38%	37%
Least deprived	176	176	352	62%	63%
Crime	283	281	564	100%	100%
Most deprived	57	57	114	20%	20%
Least deprived	226	224	450	80%	80%
Health Disability	283	281	564	100%	100%
Most deprived	50	50	100	18%	18%
Least deprived	233	231	464	82%	82%

Table3. SF12 General Health distribution baseline

Exposure variable	Poor	Fair	Good	Very good	Excellent	Total	Pearson chi2	P-value
Sample N=574	**89 (16%)**	**244 (43%)**	**198 (34%)**	**42 (7%)**	**1 (0.2%)**	**574**		
Sex	**89**	**244**	**198**	**42**	**1**	**574**	**3.16**	**0.53**
Female	45	131	94	19	1	290		
Male	44	113	104	23	0	284		
Age	**89**	**244**	**198**	**42**	**1**	**574**	**62.99**	**<0.001**
20-29	2	0	0	0	0	2		
30-39	1	1	2	0	0	4		
40-49	1	1	3	1	0	6		
50-59	12	13	6	1	0	32		
60-69	14	51	34	12	0	111		
70-79	37	91	70	18	0	216		
80-89	21	79	69	8	0	177		
90-99	1	8	14	2	1	26		
PAM13 Level	**64**	**176**	**150**	**30**	**1**	**421**	**69.34**	**<0.001**
1.Disengaged&overwhelmed	31	46	17	3	0	97		
2.Aware&struggling	14	51	34	4	0	103		
3.Taking action	12	52	58	6	0	128		
4.Maintaining behaviours& pushing further	7	27	41	17	1	93		

Table4.PAM13 distribution baseline

Exposure Variable	Level 1 (n=97) 23%	Level 2 (n=103) 24%	Level 3 (n=128) 30%	Level 4 (n=93) 22%	Total (n=421)	Pearson chi2	P-value
Sex	**97**	**103**	**128**	**93**	**421**	**3.19**	**0.36**
Female	59	56	63	48	226		
Male	38	47	65	45	195		
Age	**97**	**103**	**128**	**93**	**421**	**17.81**	**0.47**
20-29	0	0	0	0	0		
30-39	0	0	0	2	2		
40-49	2	0	2	0	4		
50-59	6	4	4	5	19		
60-69	19	18	20	22	79		
70-79	31	35	52	32	150		
80-89	35	39	41	27	142		
90-99	4	7	9	5	25		
General health	**97**	**103**	**128**	**93**	**421**	**39.80**	**<0.001**
Poor	77	65	64	34	240		
Good	20	38	64	59	181		
Mental health	**97**	**103**	**128**	**93**	**421**	**102.43**	**<0.001**
Poor	76	71	70	34	251		
Good	21	32	58	59	170		
Physical health	**97**	**103**	**128**	**93**	**421**	**36.23**	**<0.001**
Poor	92	86	95	56	329		
Good	5	17	33	37	92		

Table5. Responsiveness to intervention (*n*=288) before/after intervention

Variable	Before					After					P-Value*	% change
	Obs	Mean	Std. Err.	Std. Dev.	CI	Obs	Mean	Std. Err.	Std. Dev.	CI		
PAM13 Level	226	0.46	0.03	0.50	0.54-0.67	195	0.61	0.04	0.49	0.39-0.52	0.002	33
General health	286	0.39	0.03	0.49	0.33-0.44	288	0.45	0.03	0.50	0.39-0.51	0.125	16
Mental health	286	55.04	1.18	19.87	52.72-57.35	288	58.17	1.24	20.97	55.74-60.6	0.067	6
Physical health	286	40.48	1.32	22.40	37.87-43.09	288	44.17	1.37	23.24	41.47-46.86	0.054	9

*Analysis for equal variance

Table6.PAM13 Level as predictor (*n*=421)

SF12 Health outcome	Prob> F	R-squared	Adjusted-square	Coef.	P>t	CI
Before/after&PAM13 Level	0.001	0.0248	0.0225	-0.0732	0.001	-0.12- -0.03
General health&PAM13 Level	<0.001	0.1216	0.1195	0.32585	<0.001	0.24-0.41
Physical Health&PAM13 Level	<0.001	0.1523	0.1503	8.25915	<0.001	6.39-10.13
Mental Health&PAM13 Level	<0.001	0.1329	0.1308	6.81066	<0.001	5.14-8.48

Table7.Results multiple linear regression between SF12 and PAM13

PAM13 level		Health Outcomes	Coef.(95% CI)	
	n	1.General Health[2]	2.Mental health [3]	3.Physical health [4]
Unadjusted[1]	**421**	0.33(0.24-0.41)	6.81(5.14-8.48)	8.26(6.39-10.13)
Level 1 Disengaged & overwhelmed	97	Reference	Reference	Reference
Level 2 Aware but struggling	103	0.39(0.13-0.65)	6.74(1.65-11.83)	9.05(3.22-14.88)
Level 3 Taking action	128	0.63(0.38-0.87)	11.84(6.98-16.70)	15.89(10.33-21.44)
Level 4 Maintaining behaviour	93	1.05(0.78-1.31)	20(15.00-25.51)	24.85(18.87-30.83)
R^2	421	0.16	0.20	0.16
Adjusted R^2	421	0.15	0.18	0.15
Prob>F*		<0.001	<0.001	<0.001

1) Unadjusted Health Outcome and PAM13 Level 2)Adjusted for age 3)Adjusted for sex and age 4)Adjusted for sex
*F-statistic test decides whether the model as a whole is significant, compared to no model at all. Test H0=model is not significant rejected if the Prob>F is less than 0.05.

REFERENCES

1. Deeny S, Ruth Thorlby, Adam Steventon. Briefing: Reducing emergency admissions: unlocking the potential of people to better manage their long-term conditions [Internet]. 2018 Aug. Available from: https://www.health.org.uk/sites/default/files/Reducing-Emergency-Admissions-long-term-conditions-briefing.pdf

2. NHS England. A&E Attendances and Emergency Admissions 2017-18 [Internet]. Available from: www.england.nhs.uk/statistics/statistical-work-areas/ae-waiting-times-and-activity/ae-attendances-and-emergency-admissions-2017-18

3. NHS England. The NHS Long Term Plan [Internet]. 2019 Jan. Available from: https://www.longtermplan.nhs.uk

4. Kivelä K, Elo S, Kyngäs H, Kääriäinen M. The effects of health coaching on adult patients with chronic diseases: A systematic review. Patient Educ Couns. 2014 Nov;97(2):147–57.

5. Coulter A, Entwistle VA, Eccles A, Ryan S, Shepperd S, Perera R. Personalised care planning for adults with chronic or long- term health conditions (Review) [Internet]. The Cochrane Collaboration.; 2015. Available from: http://aura.abdn.ac.uk/bitstream/handle/2164/5609/CD010523.pdf?sequence=1

6. Panagioti M, Reeves D, Meacock R, Parkinson B, Lovell K, Hann M, et al. Is telephone health coaching a useful population health strategy for supporting older people with multimorbidity? An evaluation of reach, effectiveness and cost-effectiveness using a 'trial within a cohort.' BMC Med. 2018 May 30;16(1):80.

7. Artetxe A, Beristain A, Graña M. Predictive models for hospital readmission risk: A systematic review of methods. Comput Methods Programs Biomed. 2018 Oct;164:49–64.

8. Steele AJ, Denaxas SC, Shah AD, Hemingway H, Luscombe NM. Machine learning models in electronic health records can outperform conventional survival models for predicting patient mortality in coronary artery disease. Singh TR, editor. PLOS ONE. 2018 Aug 31;13(8):e0202344.

9. Ostrovsky A, O'Connor L, Marshall O, Angelo A, Barrett K, Majeski E, et al. Predicting 30- to 120-Day Readmission Risk among Medicare Fee-for-Service Patients Using Nonmedical Workers and Mobile Technology. Perspect Health Inf Manag. 2016;13:1e.

10. Luo W, Nguyen T, Nichols M, Tran T, Rana S, Gupta S, et al. Is Demography Destiny? Application of Machine Learning Techniques to Accurately Predict Population Health Outcomes from a Minimal Demographic Dataset. Dowd JB, editor. PLOS ONE. 2015 May 4;10(5):e0125602.

11. Smith SM, Wallace E, O'Dowd T, Fortin M. Interventions for improving outcomes in patients with multimorbidity in primary care and community settings. Cochrane Effective Practice and Organisation of Care Group, editor. Cochrane Database Syst Rev [Internet]. 2016 Mar 15 [cited 2019 Jan 28]; Available from: http://doi.wiley.com/10.1002/14651858.CD006560.pub3

12. Hutchison AJ, Breckon JD. A review of telephone coaching services for people with long-term conditions. J Telemed Telecare. 2011 Dec;17(8):451–8.

13. McLean S, Protti D, Sheikh A. Telehealthcare for long term conditions. BMJ. 2011 Feb 3;342(feb03 2):d120–d120.

14. Olsen JM, Nesbitt BJ. Health Coaching to Improve Healthy Lifestyle Behaviors: An Integrative Review. Am J Health Promot. 2010 Sep;25(1):e1–12.

15. Dennis SM, Harris M, Lloyd J, Powell Davies G, Faruqi N, Zwar N. Do people with existing chronic conditions benefit from telephone coaching? A rapid review. Aust Health Rev. 2013;37(3):381.

16. Brady TJ, Murphy L, O'Colmain BJ, Beauchesne D, Daniels B, Greenberg M, et al. A Meta-Analysis of Health Status, Health Behaviors, and Health Care Utilization Outcomes of the Chronic Disease Self-Management Program. Prev Chronic Dis [Internet]. 2013 Jan 17 [cited 2019 Jan 31];10. Available from: http://www.cdc.gov/pcd/issues/2013/12_0112.htm

17. Blakemore A, Hann M, Howells K, Panagioti M, Sidaway M, Reeves D, et al. Patient activation in older people with long-term conditions and multimorbidity: correlates and change in a cohort study in the United Kingdom. BMC Health Serv Res [Internet]. 2016 Dec [cited 2019 Jan 29];16(1). Available from: http://bmchealthservres.biomedcentral.com/articles/10.1186/s12913-016-1843-2

18. Wennberg DE, Marr A, Lang L, O'Malley S, Bennett G. A Randomized Trial of a Telephone Care-Management Strategy. N Engl J Med. 2010 Sep 23;363(13):1245–55.

19. Härter M, Dirmaier J, Dwinger S, Kriston L, Herbarth L, Siegmund-Schultze E, et al. Effectiveness of Telephone-Based Health Coaching for Patients with Chronic Conditions: A Randomised Controlled Trial. Fukumoto Y, editor. PLOS ONE. 2016 Sep 15;11(9):e0161269.

20. Wolever RQ, Dreusicke M, Fikkan J, Hawkins TV, Yeung S, Wakefield J, et al. Integrative Health Coaching for Patients With Type 2 Diabetes. Diabetes Educ. 2010 Jul;36(4):629–39.

21. Grady PA, Gough LL. Self-Management: A Comprehensive Approach to Management of Chronic Conditions. Am J Public Health. 2014 Aug;104(8):e25–31.

22. Edgren G, Anderson J, Dolk A, Torgerson J, Nyberg S, Skau T, et al. A case management intervention targeted to reduce healthcare consumption for frequent Emergency Department visitors: results from an adaptive randomized trial. Eur J Emerg Med. 2016 Oct;23(5):344–50.

23. Sjöbom P, Andrew Philips, Fiona Bell, Sofia Landström, Joachim Werr. Transferability of innovative care models: a large scale telephone based health coaching RCT. In Dublin: Health Navigator; 2017 [cited 2019 Feb 6]. Available from: http://www.health-navigator.co.uk/publications/transferability-of-innovative-care-models-a-large-scale-telephone-based-health-coaching-rct/

24. Hibbard JH, Mahoney ER, Stockard J, Tusler M. Development and Testing of a Short Form of the Patient Activation Measure. Health Serv Res. 2005 Dec;40(6p1):1918–30.

25. Insignia Health. I would like to use the PAM® survey in a research study... [Internet]. 2019 [cited 2019 Jun 20]. Available from: https://www.insigniahealth.com/research/research-licenses

26. Rademakers J, Nijman J, van der Hoek L, Heijmans M, Rijken M. Measuring patient activation in the Netherlands: translation and validation of the American short form Patient Activation Measure (PAM13). BMC Public Health. 2012 Dec;12(1):577.

27. Moljord IEO, Lara-Cabrera ML, Perestelo-Pérez L, Rivero-Santana A, Eriksen L, Linaker OM. Psychometric properties of the Patient Activation Measure-13 among out-patients waiting for mental health treatment: A validation study in Norway. Patient Educ Couns. 2015 Nov;98(11):1410–7.

28. Melville MR. Quality of life assessment using the short form 12 questionnaire is as reliable and sensitive as the short form 36 in distinguishing symptom severity in myocardial infarction survivors. Heart. 2003 Dec 1;89(12):1445–6.

29. Smith T, Michael Noble, Stefan Noble, Gemma Wright, David McLennan, Emma Plunkett. The English Indices of Deprivation 2015, research report [Internet]. Department for Communities and Local Government; 2015 Sep [cited 2016 Jun 20] p. 94. Available from: https://www.gov.uk/government/statistics/english-indices-of-deprivation-2015

30. J. Cohen. Statistical power analysis for the behavioral sciences. 2nd ed. Lawrence Erlbaum Associates, Hillsdale, NJ;

31. How to Score SF-12 Health Survey Items [Internet]. SF-12v2 - kwaliteit van leven. [cited 2019 Jun 18]. Available from: https://www.anesthesiologie.nl/uploads/files/PR_2_-_Toelichting_SF-12v2.pdf

32. John E WareM. A. KosinskiS. D. Keller. SF-12: How to Score the SF-12 Physical and Mental Health Summary Scales 1996 [Internet]. 2 Dec, 1995. Boston, Massachusetts: The Health Institute, New England Medical Center; 1998 [cited 2019 Jun 18]. 85 p. Available from: https://www.researchgate.net/publication/242636950_SF-12_How_to_Score_the_SF-12_Physical_and_Mental_Health_Summary_Scales

33. Blakemore A, Hann M, Howells K, Panagioti M, Sidaway M, Reeves D, et al. Patient activation in older people with long-term conditions and multimorbidity: correlates and change in a cohort study in the United Kingdom. BMC Health Serv Res. 2016 Dec;16(1):582.

34. Pifarre, M, K. Frost. PAM: Measuring Patient Activation in South Eastern Sydney [Internet]. Australia, Sydney: South Eastern Sydney Medicare Local; 2015 Apr [cited 2019 Jul 2]. Available from: https://www.seslhd.health.nsw.gov.au/sites/default/files/migration/CDM/documents/reports/2015/ACI%20PAM%20Project%20Final%20Report%20April%202015.pdf

35. Sterne JAC, Savović J, Page MJ, Elbers RG, Blencowe NS, Boutron I, Cates CJ, Cheng H-Y, Corbett MS, Eldridge SM, Hernán MA, Hopewell S, Hróbjartsson A, Junqueira DR, Jüni P, Kirkham JJ, Lasserson T, Li T, McAleenan A, Reeves BC, Shepperd S, Shrier I, Stewart LA, Tilling K, White IR, Whiting PF, Higgins JPT. RoB 2: a revised tool for assessing risk of bias in randomised trials. [Internet]. BMJ (in press); 2019 Jul. Available from: https://sites.google.com/site/riskofbiastool/welcome/rob-2-0-tool?authuser=0

36. Bower P, Reeves D, Sutton M, Lovell K, Blakemore A, Hann M, et al. Improving care for older people with long-term conditions and social care needs in Salford: the CLASSIC mixed-methods study, including RCT. Health Serv Deliv Res. 2018 Aug;6(31):1–188.

37. Higgins JPT SG. Cochrane Handbook for Systematic Reviews of Interventions [Internet]. 5.1.0. 2011 [cited 2018 Jan 3]. Available from: http://handbook-5-1.cochrane.org

APPENDICES

Appendix 1. SF12 and PAM13 Questionnaires

1.1 Medical Outcomes 12-Item Short Form Survey Instrument (SF-12)

1. In general, would you say your health is:

1 - Excellent
2 - Very good
3 - Good
4 - Fair
5 – Poor

2. The following questions are about activities you might do during a typical day. Does your health now limit you in these activities? If so, how much?

	Yes, limited a lot	Yes, limited a little	No, not limited at all
a. The kinds or amounts of vigorous activities you can do, like lifting heavy objects, running or participating in strenuous sports	1	2	3
b. The kinds or amounts of moderate activities you can do, like moving a table, carrying groceries, or bowling	1	2	3
c. Walking uphill or climbing a few flights of stairs	1	2	3
d. Bending, lifting, or stooping	1	2	3
e. Walking one block	1	2	3
f. Eating, dressing, bathing, or using the toilet	1	2	3

3. How much bodily pain have you had during the past 4 weeks:

1 - None
2 - Very mild
3 - Mild
4 - Moderate
5 - Severe
6 - Very Severe

4. Does your health keep you from working at a job, doing work around the house, or going to school?

1 - YES, for more than 3 months
2 - YES, for 3 months or less
3 – NO

5. Have you been unable to do certain kinds or amounts of work, housework, or schoolwork because of your health?

1 - YES, for more than 3 months
2 - YES, for 3 months or less
3 – NO

For each of the following questions, please mark the circle for the one answer that comes closest to the way you have been feeling during the past month.

	All of the time	Most of the time	A good bit of the time	Some of the time	A little of the time	None of the time
6. How much of the time, during the past month, has your health limited your social activities (like visiting with friends or close relatives)?	1	2	3	4	5	6
7. How much of the time, during the past month, have you been a very nervous person?	1	2	3	4	5	6
8. During the past month, how much of the time have you felt calm and peaceful?	1	2	3	4	5	6
9. How much of the time, during the past month, have you felt downhearted and blue?	1	2	3	4	5	6
10. During the past month, how much of the time have you been a happy person?	1	2	3	4	5	6
11. How often, during the past month, have you felt so down in the dumps that nothing could cheer you up?	1	2	3	4	5	6

12. Please mark the circle that best describes whether each of the following statements is true or false for you.

	Definitely true	Mostly true	Not sure	Mostly false	Definitely false
a. I am somewhat ill	1	2	3	4	5
b. I am as healthy as anybody I know	1	2	3	4	5
c. My health is excellent	1	2	3	4	5
d. I have been feeling bad lately	1	2	3	4	5

1.2. Patient Activation Measure (PAM-13)

Patient Activation Measure (PAM-13 describes the knowledge, skills and confidence a person has in managing their own health and care.

Below are some statements that people sometimes make when they talk about their health. Please indicate how much you agree or disagree with each statement as it applies to you personally by circling your answer. There are no right or wrong answers, just what is true for you. If the statement does not to apply you, circle N/A

1	I am the person who is responsible for taking care of my health.	Disagree Strongly	Disagree	Agree	Agree Strongly	NA
2	Taking an active role in my own health care is the most important thing that affects my health.	Disagree Strongly	Disagree	Agree	Agree Strongly	NA
3	I am confident I can help prevent or reduce problems associated with my health.	Disagree Strongly	Disagree	Agree	Agree Strongly	NA
4	I know what each of my prescribed medications do.	Disagree Strongly	Disagree	Agree	Agree Strongly	NA
5	I am confident that I can tell whether I need to go to the doctor or whether I can take care of a health problem myself.	Disagree Strongly	Disagree	Agree	Agree Strongly	NA
6	I am confident that I can tell a doctor or nurse concerns I have even when he or she does not ask.	Disagree Strongly	Disagree	Agree	Agree Strongly	NA
7	I am confident that I can carry out medical treatments I may need to do at home.	Disagree Strongly	Disagree	Agree	Agree Strongly	NA
8	I understand my health problems and what causes them.	Disagree Strongly	Disagree	Agree	Agree Strongly	NA
9	I know what treatments are available for my health problems.	Disagree Strongly	Disagree	Agree	Agree Strongly	NA
10	I have been able to maintain lifestyle changes, like healthy eating or exercising.	Disagree Strongly	Disagree	Agree	Agree Strongly	NA
11	I know how to prevent problems with my health.	Disagree Strongly	Disagree	Agree	Agree Strongly	NA
12	I am confident I can work out solutions when new problems arise with my health.	Disagree Strongly	Disagree	Agree	Agree Strongly	NA
13	I am confident that I can maintain lifestyle changes, like healthy eating and exercising, even during times of stress.	Disagree Strongly	Disagree	Agree	Agree Strongly	NA

Appendix 2.Supplementary Tables and Figures

2.2. Excluded vs included group differences, missing data n=591

Variable	Patients excluded (*N* = 303)				Patients included (*N* = 288)				Differences N=591		
	n	**Mean**	**SD**	**95% CI**	**n**	**Mean**	**SD**	**95% CI**	**Mean**	**95% CI**	**P-Value**
Sex	303	1.76	0.80	1.67-1.85	288	2.50	0.77	2.44-2.55	-0.737	-0.85-0.63	<0.001
Age	257	74.77	11.72	73.33-76.21	288	74.79	10.80	73.54-76.04	-0.018	-1.91-1.88	0.9853
Education	251	6.79	2.69	6.46-7.13	281	7.29	2.62	6.99-7.59	-0.499	-0.94-0.05	0.02
Crime	251	7.51	2.68	7.18-7.84	281	7.94	2.43	7.66-8.23	-0.433	-0.87-0.002	0.05

	Missing	Total	Percent Missing
Age	46	591	7.78
Deprivation	59	591	9.98
Income	59	591	9.98
Employment	59	591	9.98
Education	59	591	9.98
Health disability	59	591	9.98
Crime	59	591	9.98
Housing	59	591	9.98
Environment	59	591	9.98
Sex	0	591	0
Cohort	0	591	0

2.3.Mean ± SD for continuous variables for categorical variables outcomes age, sex, IMD socioeconomic score.

Exposure variable	After (n=288)	Before (n=286)	Total (n=574)	Pearson chi2	p-value	Mean	Std. Dev.	Min	Max
Sex	288 50%	286 50%	574	0.12	1.000	74.78	10.80	27	98
Age	288 50%	286 50%	574	0.01	1.000	74.78	10.80	27	98
SF General health binary	288 50%	286 50%	574	2.36	0.125	37.09	24.95	0	100
SF General health	288 50%	286 50%	574	6.70	0.152	2.48	1.00	1	5
SF Mental health composite	**288** 50%	**286** 50%	**574**	**79.92**	**0.005**	**56.61**	**20.47**	**0.00**	**100.00**
SF Physical health composite	288 50%	286 50%	574	81.02	0.218	42.33	22.88	0.00	98.13
PAM13 Level	**195** **34%**	**226** **34%**	**421** **68%**	**12.64**	**0.005**	**2.52**	**1.07**	**1**	**4**
Deprivation	283 49%	281 49%	564 99%	0.04	1.000	7.80	2.20	2	10
Income	283 49%	281 49%	564 99%	0.02	1.000	7.72	2.24		
Employment	283 49%	281 49%	564 99%	0.04	1.000	7.63	2.18	2	10
Education	283 49%	281 49%	564 99%	0.03	1.000	7.30	2.55	1	10
Housing	283 49%	281 49%	564 99%	0.03	1.000	6.06	2.82	1	10
Environment	283 49%	281 49%	564 99%	0.03	1.000	6.12	2.60	1	10

2.4 r= Pearsons correlation coefficients between scales and summary scores of patients (n=288)

	SF12 General health	SF12 Mental health	SF12 Physical health	PAM13 Level
SF12 Mental Health	0.5212	1		
	<0.001			
	574	574		
SF12 Physical Health	0.7311	0.6763	1	
	<0.001	<0.001		
	574	574	574	
PAM13 Level	0.3486	0.3645	0.3903	1
	<0.001	<0.001	<0.001	
	421	421	421	421
Employment	0.0219	0.0212	0.0301	0.099

	SF12 General health	SF12 Mental health	SF12 Physical health	PAM13 Level
	0.6044	0.6149	0.4754	0.0444
	564	564	564	413
Health disability	0.0187	0.0197	-0.0075	0.1413
	0.6572	0.641	0.8586	0.004
	564	564	564	413
Environment	-0.052	-0.0557	-0.1168	-0.1093
	0.2179	0.1864	0.0055	0.0263
	564	564	564	413
Sex	0.0427	0.1785	0.1604	0.0732
	0.3066	<0.001	0.0001	0.1339
	574	574	574	421
Age	0.1085	0.1433	0.0121	-0.0326
	0.0093	<0.001	0.7715	0.5051
	574	574	574	421

2.5 SF12 General health and PAM13 Level binary Maximum likelihood estimate of the odds ratio

Exposure variable	Odds Ratio	chi2(1)	P-value	CI
SF12 General health binary Maximum likelihood estimate of the odds ratio				
PAM13	3.073	30.37	<0.001	2.02-4.68
Age	1.173	4.28	0.0386	1.01-1.36
Environment	0.939	3.66	0.0559	0.88-1.00
Period	*1.062*	*0.03*	*0.8532*	*0.56-2.01*
Sex	*1.249*	*1.72*	*0.1897*	*0.90-1.74*
Housing	*0999*	*0.00*	*0.9650*	*0.94-1.06*
Health Disability	*0.973*	*0.48*	*0.4862*	*0.90-1.05*
Employment	*0.979*	*0.30*	*0.5839*	*0.91-1.06*
Deprivation	*0.971*	*0.56*	*0.4534*	*0.90-1.05*
Crime	*0.965*	*1.04*	*0.3077*	*0.90-1.03*
Education	*0.989*	*0.12*	*0.7331*	*0.93-1.06*
PAM13 Level binary Maximum likelihood estimate of the odds ratio				
Time Period	0.546	9.34	0.002	0.37-0.81
General Health	1.791	36.04	<0.001	1.48-2.17
Mental health outcome	1.031	40.03	<0.001	1.02-1.04
Physical health outcome	1.030	47.58	<0.001	1.02-1.04
health disability	1.122	6.39	0.011	1.03-1.23
Environment	0.928	3.63	0.056	0.86-1.00
Sex	*1.341*	*2.23*	*0.1355*	*0.91-1.97*
Age	*0.942*	*0.41*	*0.5206*	*0.78-1.13*
Housing	*0.977*	*0.42*	*0.5146*	*0.91-1.05*
Employment	*1.084*	*3.14*	*0.0766*	*0.99-1.19*
Deprivation	*1.039*	*0.71*	*0.3994*	*0.95-1.13*

* P>chi2

2.6 Association between binary measures of SF12 General health, PAM13 Level and risk factors

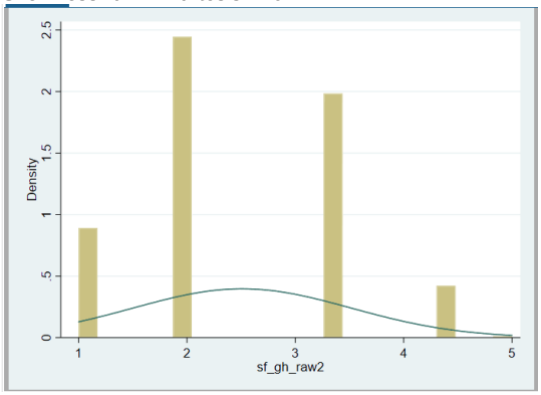
	Unadjusted[1]		Adjusted[3]				Final model[5]		
Exposure Variable	OR	95% CI	OR	95% CI	P-value[2]	% reduction[4]	OR	P-value	95% CI
Crude General Health Exposure PAM13 Level	**3.07**	**2.02-4.68**					**3.01**	**0.008**	**1.27 - 7.15**
Time period (before after intervention)	1.06	0.56-2.01	2.99	1.95-4.56	0.5653	4.21			
Environment	0.94	0.88-1.00	3.22	2.08-4.99	0.5822	-7.05			
Crime	0.97	0.90-1.03	3.16	2.05-4.85	0.3554	-4.07			
Crude Physical Health Exposure PAM13 Level	**1.03**	**1.02-1.04**					**1.03**	**0.036**	**1.00 - 1.06**
Sex			1.030	1.02-1.04	0.2989	29.82			
Age			1.030	1.02-1.04	0.9887	29.82			
Health and disability			1.031	1.02-1.04	0.8212	26.51			
Housing			1.031	1.02-1.04	0.4806	27.59			
Employment			1.031	1.02-1.04	0.9035	27.68			

	Unadjusted[1]		Adjusted[3]				Final model[5]		
Exposure Variable	OR	95% CI	OR	95% CI	P-value[2]	% reduction[4]	OR	P-value	95% CI
Environment			1.030	1.02-1.04	0.1996	29.81			

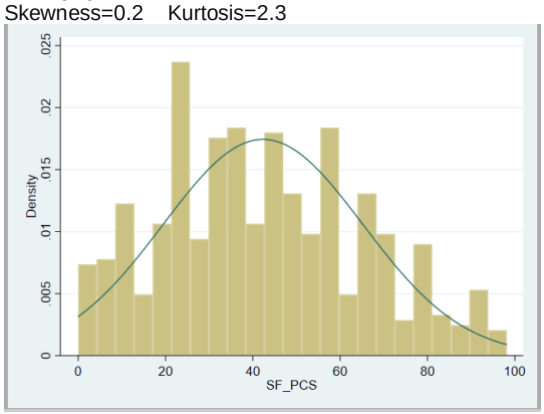
1.Maximum likelihood estimates of the odds ratio unadjusted
2. Mantel-Haenszel (MH)estimate of the odds ratio hypertension comparing education, controlling for exposure variable. Test of homogeneity of ORs (approx.)
3.Adjusted (MH Odds pooled estimate)
4.% Estimate of reduction in the effect of PAM13 Level and Health outcomes after adjustment for exposures. It was assumed that when a difference was below 5% the variable would influence the association, if over 10% the exposure variable could be a confounder.
5. Final model combined MH estimate of the odds ratio general health comes comparing PAM13 Level, controlling for period, environment and crime

2.Figure 1 Distribution skewness and kurtosis

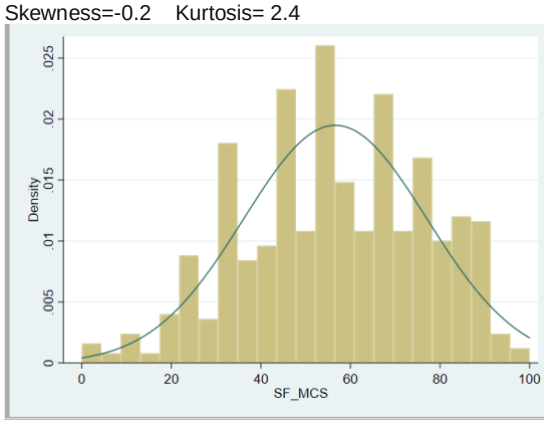
SF12 General health
Skewness=0.17 Kurtosis=2.0



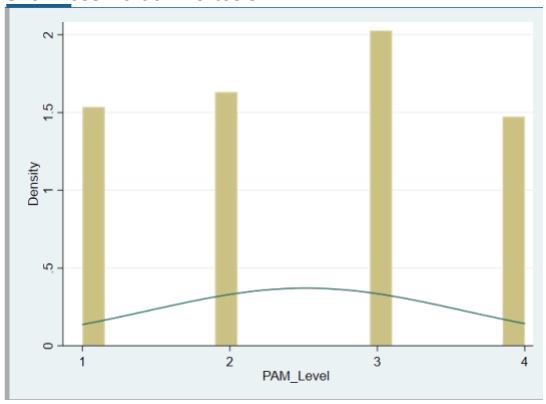
SF12 physical health
Skewness=0.2 Kurtosis=2.3



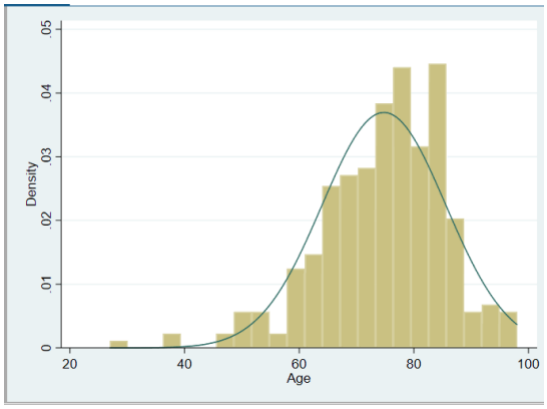
SF12 Mental health
Skewness=-0.2 Kurtosis= 2.4



PAM13 Level
Skewness=-0.06 Kurtosis=1.7



Age
Skewness=-0.8 Kurtosis=4.4



2.Figure 2. (was 16). Model adequacy