Implementation and evaluation of Exercise is Medicine in primary care clinics within a large academic health system

John Kevin Ong Dayao,1 Caroline E L Duffy,2,3,4 Amalia M Cristiano,5 Gene ‘Rusty’ Kallenberg,6 Sarah E Linke5,6,7

ABSTRACT

Objective Exercise is Medicine (EIM) is a global initiative encouraging healthcare providers to routinely assess and promote physical activity (PA) among patients. The objective of this study was to evaluate the feasibility, adoption, implementation and effectiveness of EIM from patient, clinician and healthcare staff perspectives using a combination of electronic health record (EHR), survey and interview data.

Design This study used a combination of the Practical Robust Implementation and Sustainability Model (PRISM) and the Learning Evaluation model to implement EIM. Data captured from the EHR, including Physical Activity Vital Sign (PAVS) scores, and data collected from qualitative surveys and interviews were used to evaluate the programme’s Reach, Effectiveness, Adoption, Implementation, Maintenance (RE-AIM, which is embedded within PRISM) from provider, staff and patient perspectives.

Setting Five primary care clinics within a large academic health system.

Participants A total of 24,443 patients from all participating clinics had at least one PAVS score during the study period. A total of 17 clinicians completed surveys, and 4 clinicians, 8 medical assistants and 9 patients completed interviews.

Results Implementation fidelity metrics varied widely between components and across clinics but were generally consistent over time, indicating a high degree of programme maintenance. Fidelity was highest during the first 6 months of the COVID-19 pandemic when most visits were virtual. Mean PAVS scores increased from 57.7 (95% CI: 56 to 59.4) to 95.2 (95% CI: 91.6 to 98.8) min per week at 6 months for patients not meeting PA guidelines at baseline and decreased from 253.84 (95% CI: 252 to 255.7) to 208.3 (95% CI: 204.2 to 212.4) min per week at 6 months for patients meeting PA guidelines at baseline. After EIM implementation, clinician-estimated time spent discussing PA with patients increased for 35% of providers and stayed the same for 53%.

Conclusion Overall, this study established EIM’s feasibility, adoption, implementation and maintenance in routine primary care practice within a large academic health system. From a population health perspective, EIM is a model to emulate to help primary care providers efficiently address healthy lifestyle behaviours in routine primary care visits.

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ Physical activity (PA) is effective in reducing risk factors for many health conditions including diabetes, hypertension and cardiovascular disease. Exercise is Medicine (EIM) is a global initiative that elevates PA to the level of a vital sign.

WHAT THIS STUDY ADDS

⇒ This study shows EIM, when implemented in primary care clinics within a large academic health system, is feasible and associated with improvements in PA levels among patients not meeting PA recommendations at baseline.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

⇒ Primary care clinics may add EIM into their workflow as an effective tool to support increased PA levels among their patients.

INTRODUCTION

Insufficient physical activity (PA) is a primary behavioural risk factor for numerous chronic diseases, such as hypertension, diabetes and cardiovascular disease.12 The current aerobic PA recommendation for adults is ≥150 min of moderate-intensity PA, ≥75 min of vigorous-intensity PA or a combination of the two, per week.34 Only ~50% of adults in the USA meet aerobic PA guidelines.5 Although effectiveness has been established for many PA interventions, few have been widely implemented over time with sustained outcomes.67 Implementing PA promotion in healthcare settings has been proposed as a way to leverage patient trust in healthcare providers to scale programmes that help improve a lifestyle behaviour critical to population health.8–10

Exercise is Medicine (EIM) is a global initiative supported by the American College of Sports Medicine with a goal of elevating PA to the level of a vital sign.11 12 This
framework provides clinicians with a model to routinely assess, discuss and prescribe personalised PA goals and connect patients with health coaching support and/or community-based resources to help patients achieve them. EIM has had broad global reach, and an abundance of literature describes its importance and potential impact on population-wide PA levels. Few evidence-based programmes are successfully translated into real-world settings and even fewer are delivered at scale. Implementation science (IS) and quality improvement (QI) frameworks can help to bridge the research to practice gap by guiding a systematic approach to translating, adapting, implementing and evaluating evidence-based programmes at scale.

IS is defined as the scientific study of methods to promote the systematic uptake of research findings and evidence-based practices into routine practice, and, hence, to improve the quality and effectiveness of health services. PRISM (Practical Robust Implementation and Sustainability Model), a comprehensive model for translating effective research into practice, incorporates the well-known RE-AIM (Reach, Effectiveness, Adoption, Implementation, Maintenance) model to evaluate outcomes. QI is a systematic approach to the analysis of practice performance that entails continuous efforts to reduce process variation and improve outcomes both for patients and the healthcare system. The Learning Evaluation model blends QI and implementation research methods with an emphasis on drawing systematic and transportable lessons from healthcare innovations implemented across multiple organisations in changing, real-world settings.

A unique combination of the PRISM and Learning Evaluation frameworks was used to guide the systematic adaptation, integration and evaluation of EIM in primary care clinics in a large academic health system. The purpose of this study was to evaluate the feasibility, adoption, implementation and effectiveness of the EIM initiative from patient, clinician and healthcare staff perspectives using a combination of electronic health record (EHR), survey and interview data. This systematic approach is needed to scientifically evaluate the utility and impact of EIM within the healthcare ecosystem.

METHODS
Study design
A combination of IS and QI strategies was used to promote PA assessment and intervention to the level of a vital sign. It was an uncontrolled evaluation of implementation strategies and adaptations across multiple primary care clinics. The PRISM framework was used to guide the IS strategy with a focus on evaluation, and the Learning Evaluation model was used to guide the QI strategy with focus on design and iteration. The Standards for QUAlity Improvement Reporting Excellence reporting guidelines were used in the shaping of this manuscript.

Setting
The study took place within the context of primary care clinics within the UC San Diego Health System. The clinics are located throughout San Diego County and collectively serve >120,000 patients from diverse geographical areas and backgrounds. Epic was the EHR system used to programme, document and store all patient-related activities as part of routine care. MyChart is the online patient portal connected to Epic.

Programme description
The EIM programme, implementation and adaptation processes were described in detail previously. In short, the adapted programme consisted of five core steps: (1) Physical Activity Vital Sign (PAVS) completion in the EHR; (2) automated population of the International Classification of Diseases 10 (ICD-10) code (Z71.82) on the problem and visit diagnoses lists; (3) automated reminder for primary care providers (PCPs) to document PA discussion using the EIM smartphrase, which is a tool used in the EHR to make provider documentation more efficient by populating preset phrases; (4) automated PA recommendations and resources in the patients’ After Visit Summary (AVS); and (5) automated referral for ≥1 free 15-minute health coaching telephone visit (optional, as determined by the patient’s response to a yes/no question added to the PAVS) and patient self-scheduling of this visit using the MyChart patient portal. Step 1 was required for the remaining steps to automatically happen. All steps were programmed into Epic with a local custom build that could be switched on or off for each participating primary care clinic. The final EIM workflow, following several key adaptations that were made in an iterative process to maximise efficiency, is summarised in figure 1.

For step 1, the PAVS responses were entered into Epic by either the patient via the MyChart e-check-in or by the medical assistant (MA) during the clinic’s rooming process via a local custom-built Rooming tab (Exercise); all historical PAVS data were stored and accessed in a Flowsheet.

For step 2, the Z71.82 ICD-10 code was programmed to populate with the labels of ‘Inadequate exercise’ or ‘Meeting physical activity guidelines’ depending on whether the total PAVS was less or greater than 150 min per week.

For step 3, we created a custom banner reminder and smartphrase for PCPs to document their decisions and actions around PA discussions during each patient visit in the progress note in a standardised way that enabled data collection for monitoring and analysis.

For step 4, we generated custom recommendations for weekly PA minutes based on patients’ PAVS, advising them to increase by 1 day and 5 min if they reported < 150 min and to maintain their current activity level if they reported ≥150 min as they were already meeting guidelines. The AVS also included a link to a comprehensive PA manual.
and in some clinics offered a physical copy of the manual (online supplemental file 1).

Step 5 was prompted by a yes response to the optional health coaching question in the PAVS. The AVS included a link to the MyChart portal with instructions about how to self-schedule a health coaching telephone visit, and a follow-up outreach message via MyChart was sent to patients who did not self-schedule to offer an appointment.

Data collection and retrieval
EHR data for all patients seen at participating clinics from September 2018 through January 2022 were collected as part of routine care. A de-identified and limited dataset relevant to the purpose of evaluating the implementation and effectiveness of EIM was retrieved from the backend of Epic and was sent to the principal investigator (PI) for analysis. We evaluated the programme using the RE-AIM framework, which is embedded in PRISM, as outlined below.

Reach
Reach was defined as the percentage of eligible patient visits with a PAVS recorded in the EHR during the defined study period and the demographic characteristics of these patients. The data were further stratified into patients with 1 vs >1 PAVS and patients who responded yes to the optional health coach referral and completed ≥1 health coaching visits (HCVs) versus those who did not.

Effectiveness
Effectiveness was defined in terms of PAVS changes over time. PAVS scores recorded in participating clinics during the defined study period were pulled from Epic. The first PAVS recorded for each patient was used as their own baseline (day 0). Subsequent PAVS were recorded within defined time windows of 30 (range: 14–45), 90 (range: 46–119) and 180 (range: 120–240) days post-baseline PAVS, as available. Additional health information recorded in the EHR, including diagnoses of chronic health conditions (hypertension, diabetes, obesity) and associated biometrics (body mass index (BMI) and systolic and diastolic blood pressure (SBP, DBP)), was linked to PAVS data to evaluate the effectiveness of EIM across different clinical populations and changes in biometrics over time. We did not control for other medical conditions, nor was a comorbidity index or burden measured as these were outside the scope of this study. Furthermore, changes in PAVS over time were evaluated among
patients who accepted versus declined and completed versus did not complete ≥1 HCV. All data were collected as part of routine care, with no additional data collection solely for study purposes.

Adoption
Adoption was defined as the proportion of eligible clinics that adopted EIM and the descriptive characteristics of the participating clinics that may contribute to the outcomes evaluated in the RE-AIM framework. For example, what are the demographics of the patient population served in each clinic? What are the differences in workflow between clinics that may influence EIM’s reach and implementation?

Implementation
Implementation was defined as fidelity to the five core components of EIM. Using Epic data, descriptive statistics were used to calculate the total number of eligible visits, and the proportion of visits with a PAVS, notes with EIM documentation and visits that led to a health coach referral.

Maintenance
Maintenance was defined as programme integration into routine care, as measured by fidelity to the implementation of the five core components after the active implementation period, which was 6 months for clinic 1 (pilot clinic) and 3 months for each subsequent clinic. Maintenance was explored by comparing implementation metrics achieved at 6 months with those achieved at 1 year and >1 year at each clinic.

Qualitative study
Guided by the PRISM framework, qualitative measures were used to enrich the EHR data-driven RE-AIM evaluation of EIM by capturing contextual information from key stakeholders (PCPs, MAs and patients) regarding their experiences with and perceptions about EIM. Brief online (Qualtrics) surveys were administered to PCPs and MAs, and semistructured, remote individual interviews were conducted with PCPs, MAs and patients who had recently completed an HCV. These participants were recruited via email (employees) or MyChart message (patients) and were compensated with $30 gift card incentives. All clinicians and MAs were invited, and random convenience sampling was used by inviting all patients who completed an HCV during a random week.

Data collection methods
All interviews were conducted via Zoom except for two (conducted via telephone) to accommodate participant needs. One graduate student conducted the interviews while another took detailed notes. Interviews were not recorded, and notes were de-identified to maintain participant anonymity. Average interview duration was ~20 min. The focus of the interviews was to better understand workflow, perceived barriers and benefits, areas of improvement and ideas for future directions.

Data analysis
Three reviewers independently reviewed interview notes to identify thematic results using the six-step grounded theory approach to qualitative research. We started with a purposive sample of providers and patients, followed by data collection via standardised interviews. Next, we manually generated initial codes derived from common threads identified in the data and categorised these threads into themes. Secondary or intermediate codes were then generated to synthesise the categorised themes identified in the initial coding step to assess data saturation. Finally, advanced codes from the previous synthesis were generated to develop theoretical codes from the identified theme categories. Meaning was interpreted and conclusions were drawn by the two reviewers and the PI at each step of the process, with 100% agreement on key findings.

Statistical methods
Descriptive statistics were used to calculate means, SDs and percentages of PAVS, patient demographics and biometrics. Independent t-tests and χ² tests were conducted to evaluate potential differences between patients with 1 vs ≥2 PAVS and between patients who met versus did not meet PA guidelines at the time of their first recorded PAVS. Linear mixed-effects models were used to examine changes in PAVS, BMI, SBP and DBP over time (baseline, 30, 90 and 180 days), controlling for baseline characteristics (meeting vs not meeting PA guidelines, demographics and chronic disease diagnoses) and also to examine potential differences in PAVS changes over time between patients who opted into and completed a HCV versus those who did not. Time, demographics and biometrics were included as fixed effects, and individual patients were included as random effects in the models. IBM SPSS Statistics V.28, STATA V.17 and Microsoft Excel were used for all statistical analyses, and p<0.05 was used to establish statistical significance.

Patient and public involvement
Patients were involved with all stages of this study including: programme design, where patient focus groups were conducted to elicit feedback on the PA manual; routine PCP visits where PAVS data were collected and EIM was delivered; patient surveys to elicit perspectives about EIM and patient interviews as described in this manuscript.

RESULTS
Reach
A total of 24,443 patients from all participating clinics had ≥1 PAVS recorded in Epic, including 12,938 patients with ≥2. Patients’ mean age was 51.5 (SD=17.1) years, 62% were female, 59% were non-Hispanic white, 9% had a diabetes diagnosis and 26% had a hypertension diagnosis. Mean baseline PAVS was 147.2 (SD=120.1) min per week, and ~46% of the population met PA guidelines at the time.
of their first recorded PAVS. Demographic characteristics and health condition prevalences are summarised in table 1, and comparisons are drawn between patients who: (1) had 1 vs ≥2 PAVS and (2) met versus did not meet PA guidelines at baseline. Statistically significant differences were detected between these groups, such that patients with ≥2 PAVS experienced more chronic diseases relative to those with 1 PAVS. Similarly, those not meeting PA guidelines at baseline experienced more chronic diseases relative to those meeting guidelines.

**Effectiveness**

Of the 12938 patients with ≥2 PAVS, 3226, 3026 and 3918 patients had at least one PAVS recorded within the 30-, 90-, and 180-day windows after their initial visit, respectively. These numbers were influenced by variations in visit frequency. The marginal means from the mixed-effects linear model examining change in PAVS over each of these time points, grouped according to meeting or not meeting PA guidelines at baseline, are depicted in figure 2A. From baseline to 6 months, mean PAVS of patients who met guidelines at baseline decreased from

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Demographics and health information</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total (N=24443)</td>
</tr>
<tr>
<td>Age, mean (SD)</td>
<td>51.5 (17.9)</td>
</tr>
<tr>
<td>Gender (%)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Male</td>
<td>38.2</td>
</tr>
<tr>
<td>Female</td>
<td>61.8</td>
</tr>
<tr>
<td>Race (%)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>White</td>
<td>62.2</td>
</tr>
<tr>
<td>Black</td>
<td>4.0</td>
</tr>
<tr>
<td>Asian</td>
<td>16.6</td>
</tr>
<tr>
<td>American Indian/Alaska Native</td>
<td>0.4</td>
</tr>
<tr>
<td>Native Hawaiian/Pacific Islander</td>
<td>0.6</td>
</tr>
<tr>
<td>Other or mixed race</td>
<td>14.6</td>
</tr>
<tr>
<td>Unavailable/prefer not to say</td>
<td>1.7</td>
</tr>
<tr>
<td>Ethnicity (%)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Hispanic/Latino</td>
<td>13.5</td>
</tr>
<tr>
<td>Not Hispanic Latino</td>
<td>84.6</td>
</tr>
<tr>
<td>Unavailable/prefer not to say</td>
<td>1.8</td>
</tr>
<tr>
<td>Race/ethnicity combined (%)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Non-Hispanic white</td>
<td>58.8</td>
</tr>
<tr>
<td>Non-Hispanic non-white</td>
<td>27.6</td>
</tr>
<tr>
<td>Hispanic white</td>
<td>4.5</td>
</tr>
<tr>
<td>Hispanic non-white</td>
<td>8.9</td>
</tr>
<tr>
<td>Health conditions/biometrics</td>
<td></td>
</tr>
<tr>
<td>Diabetes (%)</td>
<td>8.9</td>
</tr>
<tr>
<td>Hypertension (%)</td>
<td>26.2</td>
</tr>
<tr>
<td>Obesity status (%)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>BMI &lt;25</td>
<td>42.8</td>
</tr>
<tr>
<td>BMI 25–29.9</td>
<td>33.1</td>
</tr>
<tr>
<td>BMI ≥30</td>
<td>24.1</td>
</tr>
<tr>
<td>BMI, mean (SD)</td>
<td>26.9 (5.8)</td>
</tr>
<tr>
<td>PAVS, mean (SD)</td>
<td>147.2 (120.1)</td>
</tr>
<tr>
<td>SBP, mean (SD)</td>
<td>122.8 (15.7)</td>
</tr>
<tr>
<td>DBP, mean (SD)</td>
<td>75.5 (9.5)</td>
</tr>
</tbody>
</table>

BMI, body mass index; DBP, diastolic blood pressure; PAVS, Physical Activity Vital Sign; SBP, systolic blood pressure.
Open access

254 to 214 min per week, whereas mean PAVS of patients who did not meet guidelines increased from 58 to 95 min per week. Patients with diabetes, hypertension and/or overweight/obesity reported fewer minutes of PA per week at each time point, and the slope of the changes varied between patients with versus without these conditions (figure 2B–D).

The marginal means from the mixed-effects linear model examining changes in BMI, SBP and DBP over time according to meeting PA guidelines at baseline are depicted in figure 3. Patients not meeting guidelines at baseline had higher BMI, SBP and DBP at baseline and follow-up time points. Both groups of patients experienced significant increases in BMI and mixed changes over time in SBP and DBP (figure 3A–C, respectively).

At all time points, PAVS for patients who accepted a referral for and completed ≥1 HCV were lower compared with those who did not. Those who completed ≥2 HCVs saw a greater increase in mean PAVS over time compared with those who declined a health coaching referral, accepted but did not complete a visit or completed only one visit, but the difference between groups was not statistically significant (figure 4).

Adoption

Six clinics were invited to implement EIM, and five of the six accepted. Clinic 6 expressed interest but declined due to competing priorities. After the first four clinics had implemented EIM, clinics 2 and 5 merged (changing locations and combining providers), reducing the total number of eligible clinics to five. After the merge, EIM was implemented into the combined clinic 5, and clinic 2 was closed. The launch dates for these clinics were October 2018 (pilot clinic), October 2019, January 2020, March 2020 and July 2021, respectively. However, due to a technical glitch, clinic 5 did not have the option to use the EIM Epic tools until late September 2021, so EHR data collection was limited to 4 months.

Some differences in clinic characteristics are notable, including their locations in different geographical regions of San Diego County and their patient populations with diverse demographic and risk factor profiles. Clinics 1 and 4 are inland, whereas clinics 2, 3 and 5 are near the coast. Baseline characteristics of patients were examined using Χ² and one-way analysis of variance tests, and significant differences were noted across all baseline variables across clinics (all p<0.001). Specifically, relative to the other clinics, clinic 1 had more racially and ethnically diverse patients, and their patients had higher rates of obesity relative to the other clinics. The patient populations within clinics 2, 3 and 5 were more alike relative to clinics 1 and 4. Other notable differences between clinics included their workflow for routine primary care visits and their openness to integrating changes into their existing workflow. For example, clinics 3 and 4 asked

Figure 2 (A) PAVS over time by meeting guidelines versus not at baseline; (B) PAVS over time by having type 2 diabetes versus not; (C) PAVS over time by having hypertension versus not; (D) PAVS over time by obesity status. BMI, body mass index; PA, physical activity; PAVS, Physical Activity Vital Sign.
Figure 3  (A) Body mass index (BMI) over time by baseline PAVS—meeting guidelines versus not meeting guidelines; (B) systolic blood pressure (SBP) over time by baseline PAVS—meeting guidelines versus not meeting guidelines; (C) diastolic blood pressure (DBP) over time by baseline PAVS—meeting guidelines versus not meeting guidelines. PA, physical activity; PAVS, Physical Activity Vital Sign.

Figure 4  PAVS over time according to health coaching question (HCQ) and visit (HCV) status. PA, physical activity; PAVS, Physical Activity Vital Sign.
Dayao JKO, et al. Fam Med Com Health 2024;12:e002608. doi:10.1136/fmch-2023-002608

PAVS questions aloud during the rooming process versus relying on e-check-in or paper surveys. Because of the time required to ask patients questions aloud, the PAVS was more likely to be skipped, resulting in lower EIM rates in those clinics relative to the others.

Implementation

Table 2 summarises the key implementation metrics for each of the five clinics. Notable differences between clinics were apparent in each of the components reflected in the implementation metrics. For example, PAVS completion rates and problem list records of sufficient or insufficient PA were higher in clinics 1, 2 and 5 relative to clinics 3 and 4; the .EIM smartphrase was used more in clinics 1 and 4 relative to clinics 2, 3 and 5; and health coaching referral rates were higher among all other clinics relative to clinic 1.

We examined how these metrics varied during the first 6 months of the COVID-19 pandemic given the reduction in preventive care visits and a rapid shift to virtual care for non-acute visits at that time. As reported in table 2, fidelity to core components of EIM, particularly the proportion of visits with a recorded PAVS, increased substantially during that time.

Maintenance

Fidelity to implementation metrics at and beyond the active implementation period (6 months for clinic 1, 3 months for all other clinics) within each clinic are summarised in table 2. Within each clinic, implementation fidelity was relatively stable over time with the exception of .EIM documentation smartphrase usage, which decreased over time in most clinics. Clinic 5 implemented EIM 4 months prior to the end of data collection, limiting our ability to evaluate maintenance over time in that clinic. However, we observed that the percentages of all metrics increased each month during the first 4 months, suggesting that the programme was being absorbed as intended and indicating a high likelihood of maintenance over time.

Qualitative study

Survey

The clinician survey (n=17) found that all were somewhat (35%) or extremely (65%) comfortable discussing exercise with patients, and a majority reported that neither their comfort (76%) nor confidence (82%) levels in discussing PA changed since EIM implementation. Most reported discussing exercise with patients at non-acute visits frequently (76%) or always (6%), and 53% reported their frequency of PA discussions with patients had increased since EIM implementation. Clinicians reported discussing exercise briefly during visits, with 23% spending <1 min, 41% spending 1–3 min and 35% spending 3–6 min. Time spent discussing exercise increased for 35% of clinicians and stayed the same for 53% after EIM implementation. Reported barriers to discussing exercise with patients included lack of time (76%) and training in how to do it well (23%), belief that it will not make a difference (18%), lack of resources to

Table 2  EIM implementation metrics through January 2022

<table>
<thead>
<tr>
<th>Clinic</th>
<th>Length of time</th>
<th>Eligible visits</th>
<th>PAVS completed</th>
<th>PA dx on problem list</th>
<th>.EIM documentation</th>
<th>Health coaching referrals</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 (FM)</td>
<td>3 years 4 months</td>
<td>37328</td>
<td>54%</td>
<td>77%</td>
<td>35%</td>
<td>4%</td>
</tr>
<tr>
<td>0-12 months</td>
<td>12304</td>
<td>47%</td>
<td>59%</td>
<td>39%</td>
<td>1%</td>
<td></td>
</tr>
<tr>
<td>0-6 months</td>
<td>6344</td>
<td>41%</td>
<td>51%</td>
<td>49%</td>
<td>0%</td>
<td></td>
</tr>
<tr>
<td>March–Aug 2020*</td>
<td>3173</td>
<td>71%</td>
<td>86%</td>
<td>25%</td>
<td>5%</td>
<td></td>
</tr>
<tr>
<td>2 (IM)</td>
<td>1 year 9 months</td>
<td>12335</td>
<td>52%</td>
<td>70%</td>
<td>12%</td>
<td>9%</td>
</tr>
<tr>
<td>0-12 months</td>
<td>6641</td>
<td>56%</td>
<td>63%</td>
<td>16%</td>
<td>10%</td>
<td></td>
</tr>
<tr>
<td>0-6 months</td>
<td>3315</td>
<td>44%</td>
<td>47%</td>
<td>28%</td>
<td>11%</td>
<td></td>
</tr>
<tr>
<td>March–Aug 2020*</td>
<td>2289</td>
<td>71%</td>
<td>74%</td>
<td>14%</td>
<td>9%</td>
<td></td>
</tr>
<tr>
<td>3 (FM)</td>
<td>2 years</td>
<td>26576</td>
<td>33%</td>
<td>33%</td>
<td>15%</td>
<td>8%</td>
</tr>
<tr>
<td>0-12 months</td>
<td>12591</td>
<td>33%</td>
<td>28%</td>
<td>18%</td>
<td>8%</td>
<td></td>
</tr>
<tr>
<td>0-6 months</td>
<td>6122</td>
<td>42%</td>
<td>24%</td>
<td>17%</td>
<td>8%</td>
<td></td>
</tr>
<tr>
<td>March–Aug 2020*</td>
<td>3933</td>
<td>55%</td>
<td>31%</td>
<td>13%</td>
<td>7%</td>
<td></td>
</tr>
<tr>
<td>4 (FM)</td>
<td>1 year 10 months</td>
<td>3695</td>
<td>39%</td>
<td>46%</td>
<td>23%</td>
<td>9%</td>
</tr>
<tr>
<td>0-12 months</td>
<td>15974</td>
<td>39%</td>
<td>40%</td>
<td>24%</td>
<td>11%</td>
<td></td>
</tr>
<tr>
<td>0-6 months</td>
<td>7959</td>
<td>47%</td>
<td>39%</td>
<td>24%</td>
<td>12%</td>
<td></td>
</tr>
<tr>
<td>March–Aug 2020*</td>
<td>5161</td>
<td>52%</td>
<td>36%</td>
<td>23%</td>
<td>13%</td>
<td></td>
</tr>
<tr>
<td>5 (IM)</td>
<td>4 months</td>
<td>9832</td>
<td>51%</td>
<td>45%</td>
<td>10%</td>
<td>10%</td>
</tr>
<tr>
<td>Grand total</td>
<td>116766</td>
<td>52149 (46%)</td>
<td>64699 (54%)</td>
<td>12337 (24%)</td>
<td>3659 (7%)</td>
<td></td>
</tr>
</tbody>
</table>

*First 6 months of the COVID-19 pandemic.

dx, diagnosis; EIM, exercise is medicine; FM, family medicine; IM, internal medicine; PA, physical activity; PAVS, physical activity vital sign.
provide patients (18%) and knowledge (12%), uncertainty about how to answer questions the patient might have (6%) and lack of motivation (6%).

Self-reported utilisation of the .EIM smartphrase was higher in this sample relative to the overall rates discovered through Epic data, suggesting that the clinicians who completed the survey may not have been representative of all clinicians in the participating clinics. Among the 17 clinicians who participated in the survey, 18% reported never using the phrase, and 23%, 12%, 18% and 29% reported using it rarely, sometimes, frequently or always, respectively. Barriers to smartphrase usage included forgetting, time constraints, preference for their own documentation and not understanding its purpose when exercise is not discussed. Challenging areas identified in the clinician survey were used to inform the questions asked during the clinician interviews. All MAs who completed the survey (n=3) frequently or always collected/verified PAVS data, and all were familiar with at least one aspect of EIM, though most (57%) thought resources and health coaching had minimal impact on patients’ PA.

**Interviews**

Key themes that emerged from the structured interviews are summarised in table 3. Key themes from patients included appreciating the ease of scheduling HCVs and the overall flexibility of EIM, a desire for greater personalisation in EIM recommendations, suggesting that the reach be as broad and inclusive as possible, and a desire for increased health coaching accountability and exercise resources accessibility.

**DISCUSSION**

Overall, this comprehensive evaluation of the systematic implementation of EIM in five primary care clinics within a large academic health system indicated that it was feasible and acceptable. Reach metrics indicated that approximately half of primary care visits collected PAVS data from patients, and effectiveness metrics indicated that patients who were not meeting guidelines at baseline significantly increased their minutes of PA over time, whereas patients who met guidelines at baseline continued to surpass them at follow-up time points despite significantly decreasing over time. Adoption was high across clinics with various characteristics. Implementation metrics were mixed, with some programme components implemented at approximately half of eligible visits and others at much lower rates. Maintenance of implementation fidelity over time showed some decline, which was more or less pronounced in each of the five clinics.

Although implementation and maintenance metrics demonstrated that most of the core EIM programme components were delivered consistently over time, room for improvement in implementation fidelity was evident. Smartphrase usage was identified as a particularly challenging area in clinician surveys and Epic data, which indicated usage in only 24% of eligible visits across clinics. These data did not align with self-reported frequency of PA discussions from both the clinician survey and previously published results of patient surveys, which indicated that two-thirds of patients reported PA discussion with their PCPs. The smartphrase was designed to increase documentation efficiency, but more work should be done to encourage clinician usage (eg, helping them to identify potential benefits, using incentives, etc) or to find alternative documentation methods that can easily track the frequency of these important discussions.

Another area flagged as having room for improvement was the percentage of visits with recorded PAVS, which varied widely between clinics, ranging from 33% to 54%. Clinics 1, 2 and 5 had higher rates of recorded PAVS relative to clinics 3 and 4. A hypothesised key reason for that difference is that clinics 1, 2 and 5 routinely encourage patients to use e-check-in, which sends patients reminders to complete previsit questionnaires (including the PAVS) directly in MyChart prior to their appointment. Clinics 3 and 4, on the other hand, primarily rely on MAs to ask check-in questions (including the PAVS) aloud and manually record them in Epic, which is more time-consuming and thus more likely to be missed. This hypothesis was supported by data from the first 6 months of the COVID-19 pandemic (March through August 2020), where most primary care visits were virtual, forcing
patients to use e-check-in for most appointments. EIM programme fidelity for PAVS completion peaked in all four clinics that had implemented EIM at that time. MA qualitative interviews also supported this hypothesis, as they reported that e-check-in is more efficient. Adopting e-check-in at the time of in-office visits would also lessen MA burden altogether.

The survey and qualitative interviews indicated consensus among clinicians and MAs around the benefit of PA assessment and discussions. However, many expressed a desire to focus efforts more narrowly, such as during preventive appointments, and the need to address banner fatigue and competing priorities. This provider feedback was in stark contrast to patient feedback, however, which indicated that patients desire more opportunities to discuss PA and feel that it should be addressed universally to ensure equity across the patient population.

Health coach self-referral rates (7% of all visits at which patients were asked) and completion of at least one HCV (10% of those who self-referred) were low at all participating clinics. This was surprising given the efficient referral and scheduling processes built in Epic and the low patient burden of completing 15-minute telephone visits scheduled at their convenience. Reasons for these lower-than-expected rates were not in the scope of this study but should be explored in future studies.

Further, analyses examining potential differences in PAVS changes over time among patients who did not have health coaching referrals and completed visits found no differences despite an abundance of literature supporting the effectiveness of PA health coaching in primary care patient populations. Future research should examine the potential impact of higher doses of health coaching relative to the low-touch approach implemented in this study for scalability. More specifically, future EIM research should study more HCVs, as a meta-analysis found that PA interventions with ≥5 contacts had a larger effect on PA levels compared with those with <5 contacts. Moreover, this is an especially important area of research in the next few years as reimbursable health coaching current procedural terminology (CPT) codes are explored and established.

This study adds to existing literature such as the EIM Greenville study, which enrolled patients in a 12-week PA programme. In contrast, our implementation efforts offered a high-reach, low-touch approach. While their study found a significant decrease in body weight and SBP in patients enrolled in their PA programme, our study saw an increase in PAVS and BMI and mixed changes in SBP and DBP over time depending on baseline PA category. This finding aligns with a meta-analysis of PA interventions in primary care settings that did not lead to a decrease in BMI. This same meta-analysis revealed a 24 min per week increase relative to controls if PA was self-reported. Although our study did not have a control group, our results were consistent with these findings in that EIM was associated with increased self-reported PA over time.

In 2015, a study found a near-even split between providers who felt that EIM was helpful versus neutral or not helpful. While that study identified use of paper referral as their main barrier, our study identified lack of time and training among MAs as main barriers, which is consistent with existing literature. Future areas of improvement may explore ways to further streamline PAVS entry and emphasising EIM documentation by encouraging providers to add the smartphrase to their note templates. Addressing time constraints remains challenging given competing priorities within the broader healthcare structure. Overall, our study suggests that the perceived benefits of EIM outweighed the perceived barriers since providers reported increasing their frequency of PA discussions.

**Strengths and limitations**

Strengths of this study include its unique approach of systematically integrating EIM into a large academic health system using a combination of IS and QI methodologies, its large, diverse and universal patient population, and its triangulation of EHR, survey and qualitative interview data. All these characteristics increase generalisability when compared with a randomised controlled trial, which is a more rigorous design but captures a smaller and less representative sample.

This study is limited primarily by its real-world evidence design; without a control group, we cannot attribute the pre/post-changes in PAVS to EIM implementation. Regression to the mean may account for the increasing or decreasing PAVS over time, depending on baseline PAVS above all other predictors. Difference in counselling between patients meeting guidelines at baseline versus not may have also influenced changes in PAVS over time. Another limitation is that we did not evaluate cost or cost-effectiveness, which would also inform the scalability of this programme. Nevertheless, this study finds promising results of improving and maintaining PA levels among patients not meeting and meeting guidelines at baseline, respectively.

Generalisability of the qualitative study’s findings may be limited given the low sample size, self-selection nature of participation and social desirability bias. However, steps were taken to minimise the impact of these limitations, such as avoiding leading questions and using a grounded theory approach to analysis, which is the gold standard for qualitative analysis. Analysis was subject to confirmation bias, but agreement was reached among all three reviewers and saturation was reached, making this less likely.

**CONCLUSION**

Overall, this systematic implementation and evaluation study found that EIM can be successfully integrated into routine primary care in a large academic health system.
An evaluation of the programme more than 3 years after its initial introduction into the system using the RE-AIM framework suggests that it was adopted by the majority of participating clinics, its core components were consistently implemented and maintained over time, it reaches patients at approximately half of all non-acute visits, and it is associated with improving and maintaining PA levels among patients not meeting and changes PA and other health indicators over time. Moreover, cost analyses should examine the potential cost-effectiveness of this low-touch, high-reach approach.

**Author affiliations**
1. University of California San Diego School of Medicine, La Jolla, California, USA
3. The University of Texas Medical Branch at Galveston School of Medicine, Galveston, Texas, USA
4. Scripps Family Medicine Residency, Scripps Mercy Hospital Chula Vista, Chula Vista, California, USA
5. Herbert Wertheim School of Public Health & Human Longevity Science, University of California San Diego, La Jolla, California, USA
6. Department of Family Medicine, University of California San Diego, La Jolla, California, USA
7. Omada Health, San Francisco, California, USA

**Acknowledgements** This study was funded by KO8HS026517 (Linke, PI). The pilot study was funded by two internal grants: a UC San Diego Department of FMHP Pilot Award (Linke, PI) and an Academic Senate Award (Linke, PI). All authors declare that they have no conflicts of interest.

**Contributors** JKOD—lead on writing original draft, lead on reviewing and editing, submitting author and corresponding author. CELD—equal contribution on writing original draft, reviewing and editing, and supporting on formal analysis. AMO—equal contribution on writing original draft, reviewing and editing, and supporting on formal analysis. GBK—supporting on conceptualisation, funding acquisition, investigation, methodology, project administration, resources, supervision, and reviewing and editing. SEL—guarantor, lead on conceptualisation, data curation, formal analysis, funding acquisition, investigation, methodology and supervision, and equal contribution on writing original draft, reviewing and editing.

**Funding** This study was funded by the Agency for Health Care Research and Quality (KO8HS026517 PI: SEL). The pilot study was funded by two internal grants: a UC San Diego Department of FMHP Pilot Award (PI: SEL) and an Academic Senate Award (PI: SEL).

**Competing interests** None declared.

**Patient and public involvement** Patients and/or the public were involved in the design, or conduct, or reporting, or dissemination plans of this research. Refer to the Methods section for further details.

**Patient consent for publication** Not required.

**Ethics approval** The study protocol was reviewed by the UC San Diego Institutional Review Board (#1800968) and certified as exempt under 45 CFR 46.101(b) categories 2 and 4, falling under the auspices of quality improvement and evidence-based practices.

**Provenance and peer review** Not commissioned; externally peer reviewed.

**Data availability statement** All data relevant to the study are included in the article or uploaded as supplemental information.

**Open access** This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/.

**ORCID iD** John Kevin Ong Dayao http://orcid.org/0009-0004-2725-9896

**REFERENCES**