BMJ Open

Combining electronic health records data from a clinical research network with registry data to examine long-term outcomes of interventions and devices: an observational cohort study

Jialin Mao [ID],[1] Michael Matheny,[2] Kim G Smolderen [ID],[3] Carlos Mena-Hurtado,[3] Art Sedrakyan,[1] Philip Goodney[4]

ABSTRACT

Objectives To assess the feasibility of assessing long-term outcomes of peripheral vascular intervention (PVI) by linking data from a clinical registry to electronic health records (EHR) data from a clinical research network.

Design Observational cohort study.

Setting Vascular Quality Initiative registry linked to INSIGHT Clinical Research Network, which aggregated EHR data from multiple institutions in New York City.

Participants Patients receiving PVI during 1 January 2013–30 November 2021 in four centres in New York City.

Primary and secondary outcome measures We examined the proportion of registry patients retained in EHR over time and predictors of EHR retention after year 1. We evaluated the implications of EHR attrition by examining amputation-free survival (AFS) in the observed data and predicted data when patients discontinued in the EHR were hypothesised to have increased risks of events than the observed average.

Results We included 1405 patients receiving PVI (age=70.8±11.2 years, 51.3% male). Among eligible patients, 75.2% were retained in EHR through year 3. Patients who aged 75 years or above (vs <65: OR 0.34, 95% CI 0.18 to 0.62), had Medicaid (vs Medicare: OR 0.41, 95% CI 0.22 to 0.79), congestive heart failure (OR 0.54, 95% CI 0.32 to 0.90), dialysis (OR 0.47, 95% CI 0.24 to 0.91) and reduced ambulation (OR 0.34, 95% CI 0.15 to 0.75) were less likely to be retained in EHR. When discontinued patients were hypothesised to have increased risks of death or amputation than observed, AFS estimates diverged from the observed data around 6–12 months.

Conclusions Studies using registry-EHR data may benefit from the timeliness of the data but may be most appropriate to focus on short-term to intermediate-term outcomes of interventions and devices. Future research is needed to investigate the value of registry-EHR linkage in facilitating short-term to intermediate-term outcome assessment following vascular interventions and advanced statistical approaches to account for non-random missing long-term data.

STRENGTHS AND LIMITATIONS OF THIS STUDY

⇒ This study investigates the feasibility, strengths and limitations of using electronic health record (EHR) data to supplement registry data to evaluate intermediate to long-term patient outcomes after peripheral vascular intervention.

⇒ Strengths of this study included the use of contemporary and aggregated registry and EHR data from multiple institutions in New York City and a rigorous effort to understand the impact of non-informative censoring due to EHR discontinuation.

⇒ Limitations of this study included the uncertainty of retention rates in other clinical areas and settings and geographic areas as well as the lack of data on the exact cause of patients' discontinuation in the EHR.

INTRODUCTION

Peripheral arterial disease (PAD) affects more than 200 million people worldwide and its prevalence continues to increase.[1] Patients with PAD are at a higher risk of death and debilitating cardiovascular events, such as limb amputation. Peripheral vascular interventions (PVIs) are endovascular techniques to relieve symptoms and reduce the likelihood of serious adverse events for patients with PAD who are significantly impaired in their daily lives and do not improve with conservative treatment.[2] It is important to monitor patient outcomes beyond the first few months after the intervention.

Real-world data are useful for studying short-term and long-term outcomes of medical interventions and devices.[3] Clinical registries often collect granular disease, procedural and device characteristics but often encounter challenges in securing outcomes information due to difficulties and expenses associated with long-term patient

BMJ Group

follow-up. Electronic health records (EHRs) routinely collect longitudinal data and can augment registries' capacity to assess patient and device outcomes. EHRs have been increasingly used in research related to various clinical topics, such as cardiovascular conditions, diabetes and dementia.[4] Many EHR studies focused on risk factor identification or risk predictions. The feasibility, utility and challenges of using EHR data to supplement registry data to assess long-term outcomes interventions of interventions and devices remain unclear.

Notably, EHR data may be subject to attrition-related issues. In the USA, EHR data sharing is often limited to a single health system or a network of several hospitals. EHR data are not routinely and widely shared between hospitals nationwide. Attritions from EHR data may arise as patients relocate or seek care outside the health system or network. A previous study using EHR from Kaiser Permanente reported that 68% of patients were retained in the EHR after 3 years.[5] Another study in Oregon reported a 3-year attrition of 33%.[6] The reduced proportion of patients retained in EHR over time has important implications for study design and analysis. A previous study using pooled EHR data for drug safety research found that using EHR data to identify outcomes had reduced sensitivity compared with using claims data.[7] Given that identifying intervention and device outcomes beyond the short term is often critical, it is imperative to understand the extent to which attrition in EHR affects the assessment of patient outcomes in this context.

The current study aimed to assess the feasibility, strengths and limitations of using EHR data to supplement registry data to evaluate intermediate to long-term patient outcomes after PVI. We sought to evaluate the retention of patients in EHR data and the implications of attrition for long-term patient outcome assessment in research of medical devices and interventions.

METHODS
Data sources
We used the Vascular Quality Initiative (VQI) registry and INSIGHT Clinical Research Network for this observational cohort study. The VQI, established by the Society for Vascular Surgery, collects demographic, clinical and procedural data on patients undergoing vascular procedures from academic and community hospitals nationwide.[8] INSIGHT is a clinical research network that gathers and harmonises data for research purposes.[9] INSIGHT contains aggregated EHR data from multiple institutions in New York City and links the EHR data to other ancillary data sources.[10] Participating institutions provide a limited dataset of their EHR data within the past 10-year block. In this study, we used the EHR component of INSIGHT. The common data model harmonises structured EHR data from all institutions, including diagnoses made and procedures performed during each encounter, lab tests and results, and medication prescriptions. Hereafter, we use 'EHR data' when referring to EHR data from the INSIGHT Clinical Research Network.

Cohort creation
We identified patients receiving PVI in four centres (Weill Cornell Medicine, New York University Langone Health, Mount Sinai, Montefiore) enrolled in both VQI and INSIGHT during 1 January 2013–30 November 2021. We applied and adapted a previously validated indirect linkage method[11] to match their procedures recorded in the registry and EHR. We focused on the first procedure for each patient. We included patients receiving two types of devices, percutaneous transluminal angioplasty (PTA) with balloons or stent placement, for unilateral femoro-popliteal occlusive disease with or without a concurrent procedure on ipsilateral infrapopliteal arteries (figure 1). We excluded patients with emergent admissions, prior lower extremity amputation and missing values for payor, smoking status, comorbidities or ambulation.

Covariates and long-term outcomes
We identified patient, disease and procedural characteristics from the registry as it collects more granular data relevant to the disease context. These covariates were chosen as they were important factors to characterise the study cohort and may potentially affect retention and long-term death or amputation. Variables examined were age, sex, race and ethnicity, primary payor, body mass index (BMI), smoking status, urgency, procedure setting (inpatient, outpatient), comorbidities (hypertension, diabetes, coronary artery disease, congestive heart failure, chronic obstructive pulmonary disease and dialysis), ambulation, limb symptom, prior lower extremity revascularisation, procedure type (balloon angioplasty, stent placement), concurrent atherectomy, concurrent infrapopliteal procedure, treatment length, and preoperative and discharge antiplatelet and statin use. EHR data were used to fill in missing values for BMI, smoking status and procedure setting whenever possible.

The main long-term outcome examined was amputation-free survival (AFS). Death and amputation were assessed separately in secondary analyses. Death was ascertained from the VQI and INSIGHT using the best information available. Deaths captured in INSIGHT were in-hospital deaths. The VQI also captured some death information during follow-up contact or the Social Security Death Index in some cases. Amputation was identified from EHR data using procedure codes.

Assessment of EHR retention
We assessed the proportion of patients retained in Insight EHR among eligible patients during each follow-up year after the index PVI. EHR retention was defined as having at least one medically attended event recorded in EHR during the specific follow-up year. Eligible patients were defined as those who did not reach the end of the study (31 December 2021) or have recorded death during that specific follow-up year. Thus, attrition in this study referred
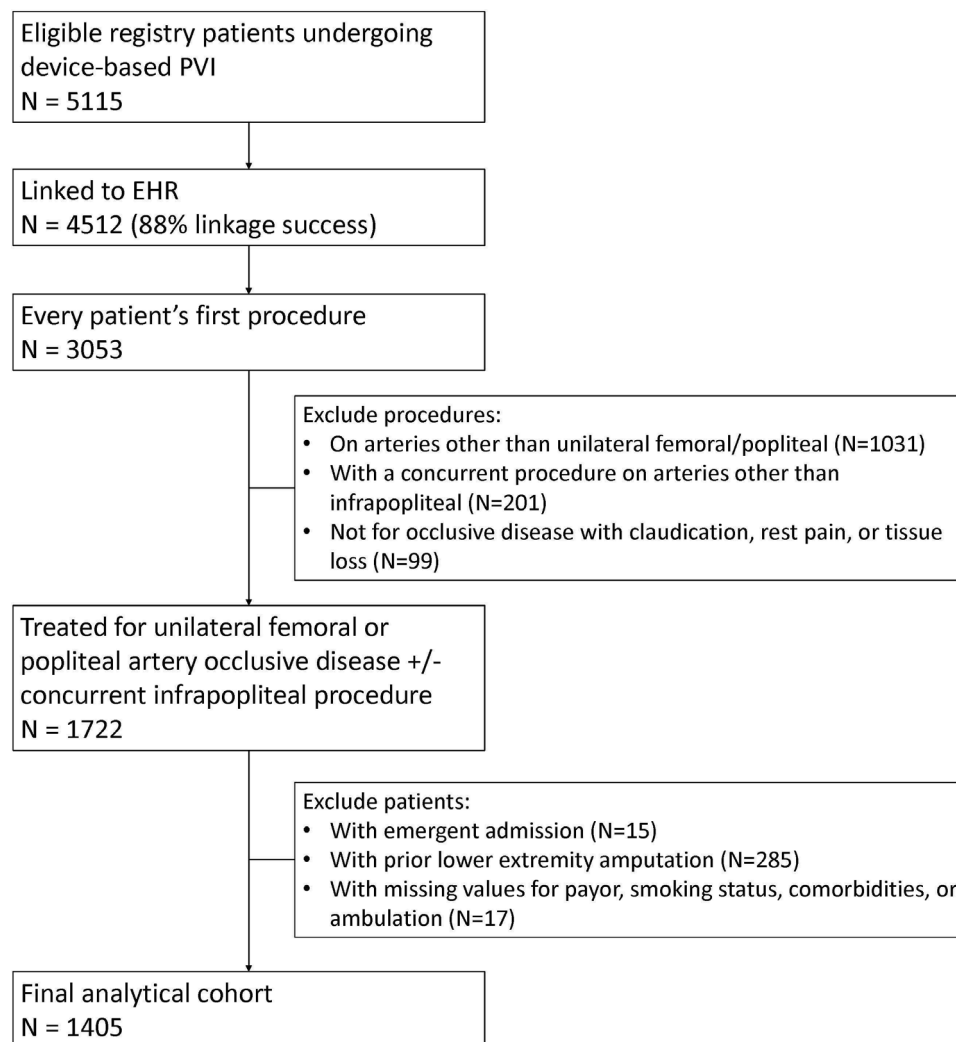
Figure 1 Flow chart demonstrating the cohort inclusion/exclusion. EHR, electronic health record; PVI, peripheral vascular intervention.

to patients who did not have any medically attended event recorded in the EHR without known death or reaching the end of the study period. We then examined the proportion of patients retained in EHR after year one by patient characteristics and predictors of EHR retention beyond 1 year using a multivariable logistic regression.

Evaluation of the implication of EHR attrition for long-term outcome assessment

To understand the implication of EHR attrition for long-term outcome assessment, we performed a hypothetical analysis by predicting events for patients who discontinued in the EHR before the end of the study. The main outcome variable was AFS. Secondary analyses were performed for death and amputation separately. We used parametric Weibull models to calculate the survival function and hazard for patients with claudication and chronic limb-threatening ischaemia (CLTI, including those with rest pain and tissue loss), respectively. The models were based on observed events and took into account known predictors of AFS, including diabetes, congestive heart failure, dialysis and ambulation. For

each patient who discontinued, starting from the time of the last recorded medically attended event, we predicted the outcome event between day T and T+1 based on the instant hazard if the patient had no events until T. We set two scenarios for instant hazard a priori: 50% increase and 100% increase compared with the observed average. These scenarios evaluated how patient outcome assessment would be impacted if patients who discontinued in the EHR were 1.5 times or twice as likely to experience events as the observed average. These hypothetical rates were chosen because previous research has shown that patients with reduced ambulation or on dialysis had 1.4–3 times higher risks of death or amputation[12–14] and these patients were more likely to have attrition in EHR. Because of the randomness in the prediction process, we performed 10 iterations of predictions for each scenario.

Based on the observed and predicted data, we examined 5-year AFS after PVI, stratified by limb symptom (claudication and CLTI), using Kaplan-Meier analyses. In the observed data, patients were censored at the end of the study or the time of the last medically attended event,

whichever was earlier. In the predicted data, patients were censored at the end of the study. We also performed a comparative analysis of AFS for patients who received PTA versus stent placement using observed and predicted data. Previous studies and current data did not demonstrate any difference in death or amputation between patients undergoing these two device-based interventions.[15–18] The procedural approach of PTA versus stent placement was also not a predictor of EHR retention in our analysis. This analysis sought to elucidate the potential impact of non-differential attrition on outcome assessment. For this analysis, we used multivariable Cox regression, adjusting for preoperative covariates. All analyses were performed by using SAS V.9.4.

Patient and public involvement

Patients or the public were not involved in the design, or conduct, or reporting, or dissemination plans of our research.

RESULTS

We identified 5115 eligible patients undergoing PVI from the VQI. Of these, 4512 (88%) were linked to the EHR. The final cohort included 1405 patients receiving PVI for unilateral femoropopliteal disease. The mean age of patients was 70.8 (±11.2) years, and 51.3% were male (table 1). 57% of patients received PTA, and 43% received stent placement. 46% of patients had intermittent claudication, and 54% had CLTI (13.5% rest pain, 40.5% tissue loss).

EHR retention

The median time from procedure to death or end of the study was 2.5 years (IQR: 1.4–4.1). Among eligible patients, 83.5% were retained in EHR through year 2, 75.2% through year 3, 69.6% through year 5 and 52.5% through year 8 (figure 2). Patients who were 75 years or above were less likely to be retained in EHR after year 1 than those aged under 65 years (OR 0.34, 95% CI 0.18 to 0.62) (table 2). Patients insured by Medicaid were less likely to be retained in EHR than those with Medicare or VA (OR 0.41, 95% CI 0.22 to 0.79). Patients who had congestive heart failure (OR 0.54, 95% CI 0.32 to 0.90) or were on dialysis (OR 0.47, 95% CI 0.24 to 0.91) were less likely to be retained in EHR than those without. Patients who needed assistance (OR 0.60, 95% CI 0.37 to 0.98) and who needed a wheelchair or were bedridden (OR 0.34, 95% CI 0.15 to 0.75) were less likely to be retained in EHR than those with full ambulation. Patients who were discharged with statin medication were more likely to be retained in EHR than those who were not (OR 1.66, 95% CI 1.03 to 2.66).

Implication of EHR attrition for long-term outcome assessment

In the observed data, AFS at 1 year was 96.5% for patients with claudication and 65.3% for those with CLTI. AFS at 5

Table 1 Characteristics of patients undergoing peripheral vascular intervention for femoral or popliteal occlusive disease in the linked data

	N (%) or mean (SD)
N total	1405
Age	
Mean (SD)	70.8 (11.2)
<65	420 (29.9%)
65–74	448 (31.9%)
75+	537 (38.2%)
Female sex	684 (48.7%)
Race and ethnicity	
Non-Hispanic white	526 (37.4%)
Non-Hispanic black	292 (20.8%)
Hispanic	369 (26.3%)
Other/unknown	218 (15.5%)
Payor	
Medicare/VA*	886 (63.1%)
Medicaid	155 (11%)
Private	364 (25.9%)
BMI	
<18.5	36 (2.6%)
18.5 to <25	482 (34.3%)
25 to <30	520 (37%)
30+	367 (26.1%)
Smoking status	
Never	491 (34.9%)
Prior	607 (43.2%)
Current	307 (21.9%)
Urgent procedure	67 (4.8%)
Outpatient procedure	814 (57.9%)
Comorbidities	
Hypertension	1263 (89.9%)
Diabetes	828 (58.9%)
Coronary artery disease	381 (27.1%)
Congestive heart failure	229 (16.3%)
COPD	161 (11.5%)
Dialysis	140 (10%)
N comorbidities	
None	60 (4.3%)
1	274 (19.5%)
2+	1071 (76.2%)
Ambulation	
Full ambulation	1049 (74.7%)
Ambulation with assistance	290 (20.6%)
Wheelchair or bedridden	66 (4.7%)
Limb symptom	
Claudication	647 (46%)
Rest pain	189 (13.5%)
Tissue loss	569 (40.5%)

Continued

Table 1 Continued

	N (%) or mean (SD)
Prior LER	414 (29.5%)
Procedure	
Balloon angioplasty	801 (57%)
Stent placement	604 (43%)
Concurrent atherectomy	418 (29.8%)
Concurrent infrapopliteal procedure	488 (34.7%)
Treatment length (cm)	
0 to <10	264 (18.8%)
10 to <20	349 (24.8%)
20+	675 (48%)
Unknown	117 (8.3%)
Preop antiplatelet	
None	373 (26.5%)
Aspirin	584 (41.6%)
P2Y12	100 (7.1%)
Dual antiplatelet	348 (24.8%)
Preop statin	1039 (74%)
Discharge antiplatelet†	
None	89 (6.3%)
Aspirin	252 (17.9%)
P2Y12	240 (17.1%)
Dual antiplatelet	817 (58.1%)
Discharge statin†	1134 (80.7%)

*A small proportion of patients had VA health benefit.
†Patients who died in hospital (<1%) did not have discharge medication and were excluded from the denominator.
BMI, body mass index; COPD, chronic obstructive pulmonary disease; LER, lower extremity revascularisation; VA, veteran administration.

years was 88.7% for patients with claudication and 46.3% for those with CLTI. Few events were identified from EHR after year 2 (N events: 252 in year 1, 38 in year 2, 28 in years 3–5). When discontinued patients were hypothesised to have a 50% increase in the risk of death or amputation than observed, AFS at 5 years was 82%–86% for patients with claudication and 34%–41% for those with CLTI. When discontinued patients were hypothesised to have a 100% increase in the risk of death or amputation than observed, AFS at 5 years was 82%–84% for patients with claudication and 31%–37% for those with CLTI. The Kaplan-Meier curves from observed and predicted data started diverging around 6–12 months (figure 3A,B). Secondary analyses of death and amputation separately showed similar results (online supplemental figure). In the comparative analysis of 5-year AFS after stent placement versus PTA (PTA as the reference group), the observed HR and 95% CI were 1.06 (0.83 to 1.35). When discontinued patients were hypothesised to have a 50% or 100% increase in the risk of death or amputation than observed, the HRs ranged between 0.94 and 1.08 (figure 3C).
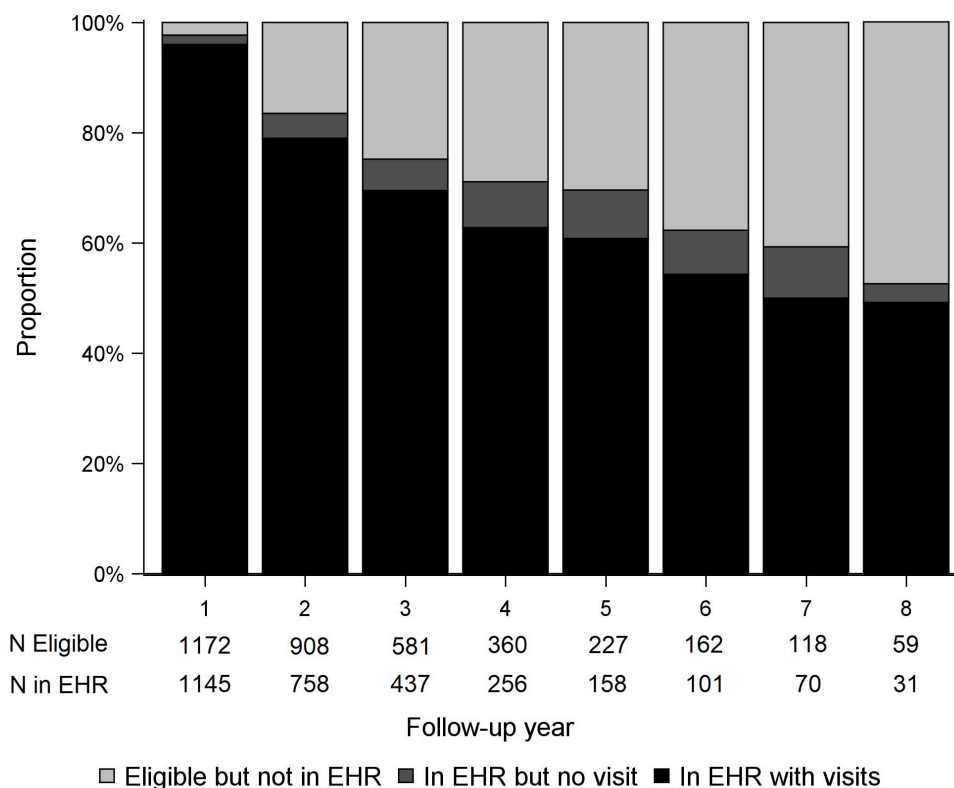
DISCUSSIONS

In this study, we used EHR data to obtain follow-ups for registry patients who underwent PVI in four academic centres. We found that 75% of patients were retained in the EHR within 3 years from the time of the index procedure. The likelihood of patients being retained in EHR differed by age, insurance, comorbidities and ambulation. We also showed that the non-random attrition in the EHR would alter absolute and comparative estimates of long-term AFS if discontinued patients were to be more likely to experience death or amputation.

Our study showed that it is feasible to use combined EHR and clinical registry data to examine outcomes of interventions and devices and the execution can be very timely. We demonstrated an 88% linkage success rate in combining the registry with EHR data. This linkage rate was similar to the linkage success rate of the VQI-Medicare linkage.[11] The 12% of procedures that cannot be linked may be due to missing or incorrectly coded procedures in the EHR. Linkage or direct mapping between EHR and registry data can enable outcome assessments for specific interventions and device attributes and designs while appropriately accounting for granular clinical details. In addition, the registry-EHR linkage was performed in 2022 and used data up to the end of 2021.

However, using EHR data to assess longer-term outcomes of medical interventions and devices can be challenging as the proportion of patients retained in EHR decreased significantly over time. We showed a 3-year attrition rate of 25%, slightly lower than the 32%–33% attrition rates reported in previous studies. This could be because we used a conservative definition of attrition and focused on patients of an older age group in this study.[5 6] But overall, these results were broadly consistent. Approximately 25%–30% of patients were not retained in the EHR beyond 3 years, which may be due to death, relocation or switching providers. These attritions may not be random and have implications for patient outcome assessment. In our study, vulnerable patients (eg, age >75 years, Medicaid insured, with reduced ambulation) were less likely to be retained in EHR. As a result, event rates may be underestimated when assuming non-informative censoring. In the data that we predicted events for discontinued patients at an increased event rate, estimates for AFS started diverging between the observed and predicted data after 6 months to 1 year. Comparative analysis may also yield different conclusions even if the exposure groups do not have differential attrition.

These results suggested that additional caution should be exercised when designing and interpreting results from long-term follow-up studies with EHR. When possible, researchers can consider using additional complementary data sources to facilitate long-term outcome assessment. For example, linking EHR to vital statistics data could help improve the reliability of long-term mortality assessment. However, even with vital statistics linkage, other long-term outcomes, such as amputation, may still be challenging to evaluate. The Vascular Implant

Follow-up year	1	2	3	4	5	6	7	8
N Eligible	1172	908	581	360	227	162	118	59
N in EHR	1145	758	437	256	158	101	70	31

☐ Eligible but not in EHR ☐ In EHR but no visit ■ In EHR with visits

Figure 2 Proportion of patients retained in EHR by year for eligible patients undergoing peripheral vascular intervention for femoropopliteal occlusive disease. EHR, electronic health record.

Surveillance and Interventional Outcomes Network is a coordinated registry network that facilitates the evaluation of vascular devices by linking the registry to other routinely collected data sources.[19] The linkage to Medicare claims data can facilitate the evaluation of long-term outcomes among Medicare beneficiaries,[20][21] but claims data are usually subject to delayed data release. Researchers need to decide, for each study context, which combination of data sources can provide the most robust and useful evidence. When additional data sources are not available and long-term outcomes are assessed from EHR data alone, advanced statistical approaches, such as Bayesian methods, may be used to estimate the impact of non-random missing data.[22] It should be noted that these methods require assumptions about data distribution. Additional research is needed to evaluate the performance of these statistical methods in handling missing data in specific disease contexts.

Whether and how the addition of EHR data to clinical registry data can facilitate the assessment of short-term to intermediate-term outcomes after vascular interventions also needs to be further investigated. EHR data have the potential to bring in lab test results, prescription data, as well as unstructured data, such as radiology reports or images.[23] Registry data collection is often limited to the perioperative time period, while EHR data capture lab results and prescription data during encounters prior to and after the index procedure. In addition, with the assistance of advanced methodologies, such as machine learning and natural language processing,[24][25]

researchers may be able to derive variables that are not otherwise available, such as vessel patency in the case of PVI follow-up. These aspects are beyond the scope of the current research and need to be evaluated in future studies.

There are limitations to this study. First, the study cohort was limited to patients treated at four academic institutions in New York City. Previous research has shown that patient retention was slightly higher in community centres.[6] The attrition rate in a larger scale EHR collaborative may also differ and needs to be assessed in future research. Additionally, we included patients undergoing PVI for PAD as a case study. EHR retention may vary for different clinical specialties and disease areas. But notably, previous studies focusing on different clinical areas (primary care and obesity) showed broadly consistent EHR retention rates.[5][6] Second, the linkage method has been validated in the prior linkage between the VQI and Medicare claims. It has not been validated in the current dataset. However, the linkage success rates were consistent. Third, our study included data from the first year of the COVID-19 pandemic. We do not expect the findings to be sensitive to any pandemic effects as the focus was on long-term attrition and the main data accumulation happened before the pandemic. However, COVID-19 may affect retention during and after the pandemic years due to possible long-lasting changes in care patterns. These impacts could not be assessed in the current study due to the limited years of data related to COVID-19. Fourth, we could not determine the exact

Table 2   Association between patient characteristics and retention in EHR after year 1 among patients undergoing peripheral vascular intervention for femoropopliteal occlusive disease

	N (%) retained in EHR	Multivariable analysis OR (95% CI)	P value
Age			
<65	240 (87.0%)	Ref	
65–74	267 (87.8%)	0.75 (0.41 to 1.36)	0.35
75+	251 (76.5%)	0.34 (0.18 to 0.62)	<0.001
Sex			
Male	398 (84.3%)	Ref	
Female	360 (82.6%)	1.14 (0.75 to 1.73)	0.55
Race/ethnicity			
Non-Hispanic white	310 (86.4%)	Ref	
Non-Hispanic black	150 (83.3%)	1.03 (0.57 to 1.85)	0.93
Hispanic	190 (81.5%)	0.80 (0.47 to 1.38)	0.42
Other/unknown	108 (79.4%)	0.58 (0.33 to 1.03)	0.06
Payor			
Medicare/VA	464 (83.0%)	Ref	
Medicaid	76 (76.8%)	0.41 (0.22 to 0.79)	0.007
Private	218 (87.2%)	0.81 (0.48 to 1.36)	0.42
BMI			
18.5 to <25	16 (76.2%)	0.95 (0.29 to 3.09)	0.93
<18.5	232 (78.4%)	Ref	
25 to <30	283 (85.0%)	1.47 (0.93 to 2.31)	0.1
30+	227 (88.0%)	1.61 (0.94 to 2.78)	0.09
Smoking status			
Never	237 (78.5%)	Ref	
Prior	362 (88.5%)	1.50 (0.93 to 2.42)	0.1
Current	159 (80.7%)	0.54 (0.31 to 0.95)	0.03
Admission status			
Elective	728 (83.5%)	Ref	
Urgent	30 (83.3%)	1.39 (0.52 to 3.69)	0.51
Setting			
Inpatient	272 (77.5%)	Ref	
Outpatient	486 (87.3%)	1.55 (0.99 to 2.41)	0.05
Hypertension			
No	100 (87.7%)	Ref	
Yes	658 (82.9%)	0.70 (0.34 to 1.45)	0.34
Diabetes			
No	341 (85.0%)	Ref	
Yes	417 (82.2%)	0.77 (0.46 to 1.29)	0.32
Coronary artery disease			
No	575 (84.9%)	Ref	
Yes	183 (79.2%)	0.63 (0.39 to 1.00)	0.05
Congestive heart failure			
No	666 (85.4%)	Ref	
Yes	92 (71.9%)	0.54 (0.32 to 0.90)	0.02

Continued

Table 2 Continued

	N (%) retained in EHR	Multivariable analysis OR (95% CI)	P value
COPD			
No	669 (83.0%)	Ref	
Yes	89 (87.3%)	1.39 (0.70 to 2.76)	0.34
Dialysis			
No	702 (84.4%)	Ref	
Yes	56 (73.7%)	0.47 (0.24 to 0.91)	0.03
N comorbidities			
0–1	207 (85.9%)	Ref	
2+	551 (82.6%)	1.06 (0.52 to 2.17)	0.87
Ambulation			
Full ambulation	610 (86.8%)	Ref	
Ambulation with assistance	126 (74.6%)	0.60 (0.37 to 0.98)	0.04
Wheelchair or bedridden	22 (61.1%)	0.34 (0.15 to 0.75)	0.008
Limb symptom			
Claudication	421 (87.9%)	Ref	
Rest pain	92 (82.1%)	1.32 (0.68 to 2.56)	0.4
Tissue loss	245 (77.3%)	1.01 (0.60 to 1.71)	0.97
Prior LER			
No	516 (82.3%)	Ref	
Yes	242 (86.1%)	1.11 (0.72 to 1.72)	0.63
Procedure			
Balloon angioplasty	418 (85.8%)	Ref	
Stent placement	340 (80.8%)	0.70 (0.46 to 1.07)	0.1
Concurrent atherectomy			
No	542 (83.0%)	Ref	
Yes	216 (84.7%)	0.99 (0.62 to 1.58)	0.97
Concurrent infrapopliteal procedure			
No	540 (85.6%)	Ref	
Yes	218 (78.7%)	0.83 (0.53 to 1.32)	0.44
Treatment length (cm)			
0-<20	330 (84.8%)	Ref	
20+	354 (82.3%)	1.10 (0.71 to 1.71)	0.68
Unknown	74 (83.1%)	1.26 (0.62 to 2.55)	0.52
Discharge antiplatelet			
No	43 (75.4%)	Ref	
Yes	715 (84.0%)	1.40 (0.67 to 2.93)	0.38
Discharge statin			
No	135 (77.1%)	Ref	
Yes	623 (85.0%)	1.66 (1.03 to 2.66)	0.04

BMI, body mass index; COPD, chronic obstructive pulmonary disease; LER, lower extremity revascularisation; VA, veteran administration .

cause of patients' discontinuation in the EHR. Future studies with qualitative components will be needed to delineate that. In addition, some of the important patient characteristics that may affect attrition and AFS, such as social determinants of health, were not collected in the data sources and cannot be assessed in this study. Lastly, our analysis of the impact of the non-random attrition on AFS estimates was constrained by the assumption we
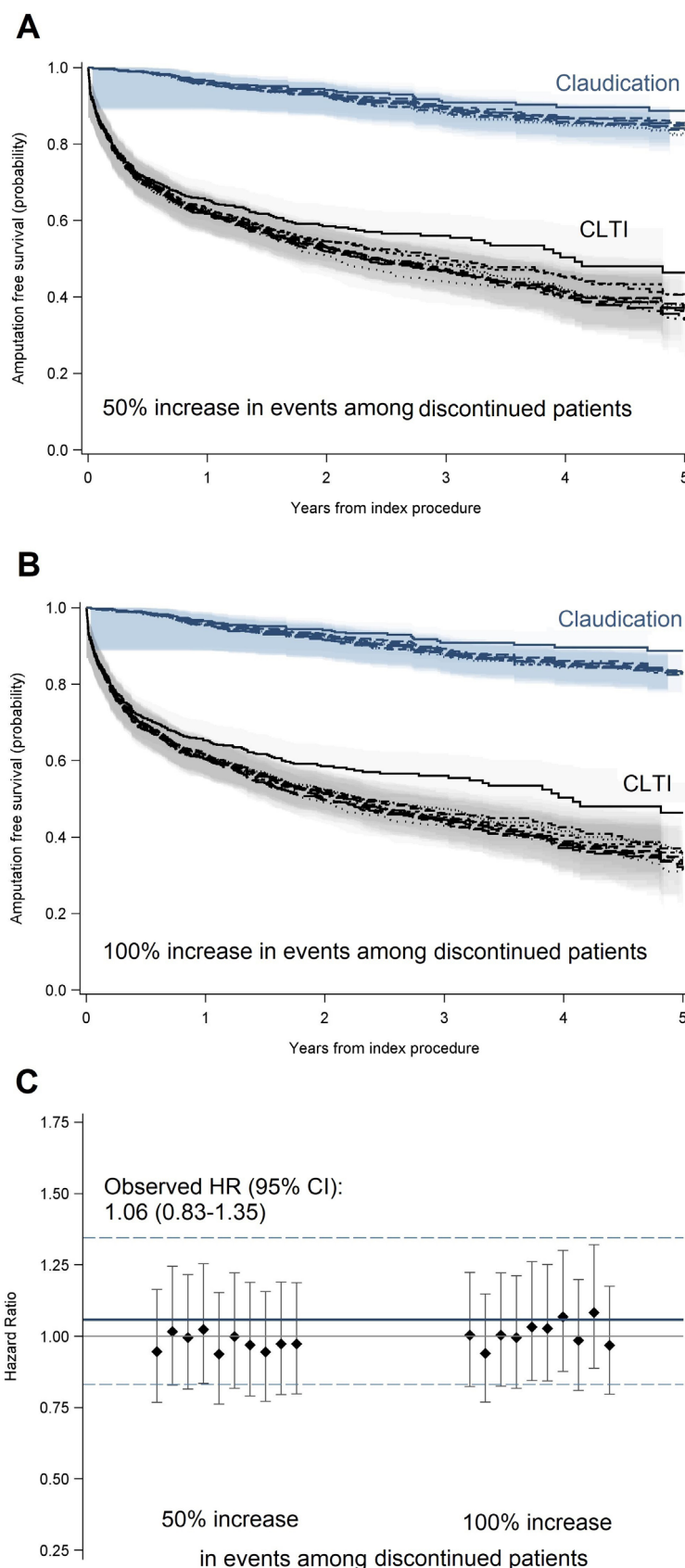
Figure 3 Hypothetical analysis of 5-year amputation-free survival with patients discontinued in the EHR at 50% and 100% increased risk of experiencing the events than average. (A, B) Kaplan-Meier analysis for the observed (solid lines) and predicted (dashed lines) events. (C) Observed (horizontal line) and predicted (vertical bars) HRs and 95% confidence bands from Cox regression comparing AFS between patients undergoing stent placement versus PTA (reference group). AFS, amputation-free survival; CTLI, chronic limb-threatening ischaemia; EHR, electronic health record; PTA, percutaneous transluminal angioplasty.

made when choosing the hypothetical rates of events for patients who discontinued in the EHR. We were not able to ascertain the 'ground truth' from the current data.

CONCLUSIONS

We presented the feasibility and limitations of combining EHR with a clinical registry to assess long-term patient outcomes following vascular interventions and device use. Research using linked registry-EHR data to assess outcomes of medical interventions and devices may benefit from the timeliness of the data but may be most appropriate to focus on short-term to intermediate-term endpoints. Future research is needed to investigate the usefulness of registry-EHR linkage in facilitating short-term to intermediate-term outcome assessment following vascular interventions and advanced statistical approaches to account for non-random missing long-term data.

Author affiliations
[1]Population Health Sciences, Joan and Sanford I Weill Medical College of Cornell University, New York, New York, USA
[2]Departments of Biomedical Informatics, Biostatistics, and Medicine, Vanderbilt University Medical Center, Nashville, Tennessee, USA
[3]Department of Internal Medicine Section of Cardiovascular Medicine, Yale School of Medicine, Yale University, New Haven, Connecticut, USA
[4]Department of Surgery Section of Vascular Surgery, Dartmouth-Hitchcock Medical Center, Lebanon, New Hampshire, USA

X Jialin Mao @JialinMao and Kim G Smolderen @kimgsmolderen

Competing interests KGS reports unrestricted research grants from Philips, Merck, Shockwave and Johnson & Johnson; she is a consultant for Optum Labs, Cook, Tegus, Twill and Abbott Vascular. CM-H reports unrestricted research grants from Abbott Vascular, Philips and Shockwave and is a consultant for Cook and Penumbra. The other authors report no competing interests.

Patient and public involvement Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

Patient consent for publication Not applicable.

Ethics approval This study was determined exempt by Brany IRB (#22-12-370-380). Informed consent was waived because this study only involved secondary analysis of existing data and it was not possible to contact patients.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement No data are available. Data used in this study were subject to data use agreements and cannot be publicly shared. Supporting documents will be made available to on contact.

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

ORCID iDs
Jialin Mao http://orcid.org/0000-0001-6671-3474
Kim G Smolderen http://orcid.org/0000-0001-6104-6254

REFERENCES

1. Fowkes FGR, Rudan D, Rudan I, *et al*. Comparison of global estimates of prevalence and risk factors for peripheral artery disease in 2000 and 2010: a systematic review and analysis. *Lancet* 2013;382:1329–40.
2. Olin JW, White CJ, Armstrong EJ, *et al*. Peripheral Artery Disease: Evolving Role of Exercise, Medical Therapy, and Endovascular Options. *J Am Coll Cardiol* 2016;67:1338–57.
3. U.S. Food and Drug Administration. Use of real-world evidence to support regulatory decision-making for medical devices. 2017. Available: https://www.fda.gov/regulatory-information/search-fda-guidance-documents/use-real-world-evidence-support-regulatory-decision-making-medical-devices [Accessed 31 Aug 2020].
4. Casey JA, Schwartz BS, Stewart WF, *et al*. Using Electronic Health Records for Population Health Research: A Review of Methods and Applications. *Annu Rev Public Health* 2016;37:61–81.
5. Koebnick C, Smith N, Huang K, *et al*. OBAYA (obesity and adverse health outcomes in young adults): feasibility of a population-based multiethnic cohort study using electronic medical records. *Popul Health Metr* 2012;10:15.
6. Huguet N, Kaufmann J, O'Malley J, *et al*. Using Electronic Health Records in Longitudinal Studies: Estimating Patient Attrition. *Med Care* 2020;58 Suppl 6 Suppl 1:S46–52.
7. Beukelman T, Chen L, Annapureddy N, *et al*. Using pooled electronic health records data to conduct pharmacoepidemiology safety studies: Challenges and lessons learned. *Pharmacoepidemiol Drug Saf* 2023;32:969–77.
8. Cronenwett JL, Kraiss LW, Cambria RP. The Society for Vascular Surgery Vascular Quality Initiative. *J Vasc Surg* 2012;55:1529–37.
9. Fleurence RL, Curtis LH, Califf RM, *et al*. Launching PCORnet, a national patient-centered clinical research network. *J Am Med Inform Assoc* 2014;21:578–82.
10. Insight Clinical Rsearch Network. 2024. Available: https://insightcrn.org/ [Accessed 02 Jul 2024].
11. Mao J, Moore KO, Columbo JA, *et al*. Validation of an indirect linkage algorithm to combine registry data with Medicare claims. *J Vasc Surg* 2022;76:266–71.
12. Miyata T, Kumamaru H, Mii S, *et al*. Prediction Models for Two Year Overall Survival and Amputation Free Survival After Revascularisation for Chronic Limb Threatening Ischaemia. *Eur J Vasc Endovasc* 2022;64:367–76.
13. Vierthaler L, Callas PW, Goodney PP, *et al*. Determinants of survival and major amputation after peripheral endovascular intervention for critical limb ischemia. *J Vasc Surg* 2015;62:655–64.
14. Kreutzburg T, Peters F, Kuchenbecker J, *et al*. Editor's Choice - The GermanVasc Score: A Pragmatic Risk Score Predicts Five Year Amputation Free Survival in Patients with Peripheral Arterial Occlusive Disease. *Eur J Vasc Endovasc* 2021;61:248–56.
15. Armstrong EJ, Saeed H, Alvandi B, *et al*. Nitinol self-expanding stents vs. balloon angioplasty for very long femoropopliteal lesions. *J Endovasc Ther* 2014;21:34–43.
16. Chalmers N, Walker PT, Belli A-M, *et al*. Randomized trial of the SMART stent versus balloon angioplasty in long superficial femoral artery lesions: the SUPER study. *Cardiovasc Intervent Radiol* 2013;36:353–61.
17. Nguyen B-N, Conrad MF, Guest JM, *et al*. Late outcomes of balloon angioplasty and angioplasty with selective stenting for superficial femoral-popliteal disease are equivalent. *J Vasc Surg* 2011;54:1051–7.
18. Schillinger M, Sabeti S, Loewe C, *et al*. Balloon angioplasty versus implantation of nitinol stents in the superficial femoral artery. *N Engl J Med* 2006;354:1879–88.
19. Tsougranis G, Eldrup-Jorgensen J, Bertges D, *et al*. The Vascular Implant Surveillance and Interventional Outcomes (VISION)

Coordinated Registry Network: An effort to advance evidence evaluation for vascular devices. *J Vasc Surg* 2020;72:2153–60.

20 Mao J, Sedrakyan A, Goodney PP, *et al*. Editor's Choice - Real World Study of Mortality After the Use of Paclitaxel Coated Devices in Peripheral Vascular Intervention. *Eur J Vasc Endovasc Surg* 2023;65:131–40.

21 Goodney P, Mao J, Columbo J, *et al*. Use of linked registry claims data for long term surveillance of devices after endovascular abdominal aortic aneurysm repair: observational surveillance study. *BMJ* 2022;379:e071452.

22 Austin PC, Escobar MD. Bayesian modeling of missing data in clinical research. *Comput Stat Data An* 2005;49:821–36.

23 Tayefi M, Ngo P, Chomutare T, *et al*. Challenges and opportunities beyond structured data in analysis of electronic health records. *Wires Comput Stat* 2021;13.

24 Xiao C, Choi E, Sun J. Opportunities and challenges in developing deep learning models using electronic health records data: a systematic review. *J Am Med Inform Assn* 2018;25:1419–28.

25 Li I, Pan J, Goldwasser J, *et al*. Neural Natural Language Processing for unstructured data in electronic health records: A review. *Comput Sci Rev* 2022;46.