

Current use and costs of electronic health records for clinical trial research: a descriptive study

Kimberly A. Mc Cord MSc, Hannah Ewald PhD MPH, Aviv Ladanie PhD, Matthias Briel MD MSc, Benjamin Speich PhD, Heiner C. Bucher MD MPH, Lars G. Hemkens MD MPH; for the RCD for RCTs initiative and the Making Randomized Trials More Affordable Group

See related article at www.cmaj.ca/lookup/doi/10.1503/cmaj.180841

Abstract

Background: Electronic health records (EHRs) may support randomized controlled trials (RCTs). We aimed to describe the current use and costs of EHRs in RCTs, with a focus on recruitment and outcome assessment.

Methods: This descriptive study was based on a PubMed search of RCTs published since 2000 that evaluated any medical intervention with the use of EHRs. Cost information was obtained from RCT investigators who used EHR infrastructures for recruitment or outcome measurement but did not explore EHR technology itself.

Results: We identified 189 RCTs, most of which (153 [81.0%]) were carried out in North America and were published recently (median year 2012 [interquartile range 2009–2014]). Seventeen RCTs (9.0%) involving a median of 732 (interquartile range 73–2513) patients explored interventions not related to EHRs, including quality improvement, screening programs, and collaborative care and disease management interventions. In these trials, EHRs were used for recruitment (14 [82%]) and outcome measurement (15 [88%]). Overall, in most of the trials (158 [83.6%]), the outcome (including many of the most patient-relevant clinical outcomes, from unscheduled hospital admission to death) was measured with the use of EHRs. The per-patient cost in the 17 EHR-supported trials varied from US\$44 to US\$2000, and total RCT costs from US\$67 750 to US\$5 026 000. In the remaining 172 RCTs (91.0%), EHRs were used as a modality of intervention.

Interpretation: Randomized controlled trials are frequently and increasingly conducted with the use of EHRs, but mainly as part of the intervention. In some trials, EHRs were used successfully to support recruitment and outcome assessment. Costs may be reduced once the data infrastructure is established.

Randomized controlled trials (RCTs) are the standard for evaluating benefits and harms of medical treatments. However, they are often time consuming and expensive to conduct, and some trials rely on strictly standardized research settings that may limit the generalizability of their results.¹ Electronic health records (EHRs) — electronic databases containing patient-level variables that are gathered during routine medical care (Appendix 1, available at www.cmajopen.ca/content/7/1/E23/suppl/DC1) — provide great potential for implementing large and pragmatic trials.^{2,3} Randomized controlled trials could be integrated directly into routine care, offering almost perfect generalizability of their results.⁴ Recently, the Patient-Centered Outcomes Research Institute awarded US\$332 million to 28 pragmatic clinical studies, many of them using EHR infrastructures and many of them integrated in routine care.⁵

Great debate regarding the potential barriers and limitations of EHR use in clinical research persists, and further details on these obstacles have been discussed elsewhere.^{3,6} Briefly, the 2 largest direct advantages of using routinely

Competing interests: Kimberly Mc Cord, Matthias Briel, Heiner Bucher and Lars Hemkens support the RCD for RCTs initiative, which aims to explore the use of routinely collected data for randomized clinical trials. Kimberly Mc Cord, Matthias Briel, Benjamin Speich and Lars Hemkens are members of the Making Randomized Trials More Affordable (MARTA) Group. No other competing interests were declared.

This article has been peer reviewed.

Correspondence to: Lars Hemkens. lars.hemkens@usb.ch

CMAJ Open 2019. DOI:10.9778/cmajo.20180096

collected data for clinical trials may be the facilitation of patient recruitment and of outcome assessment. Randomization of treatment may occur directly from the EHR during the patient's visit, maximizing recruitment rates.⁷ Recruiting patients through the EHR would allow investigators to pre-screen for eligibility before approaching the potential participant, thus enabling tailoring of the efforts toward the appropriate sample; furthermore, rapid consecutive enrolment would favour recruitment through automatic screening and selection of participants through the EHR database.⁸ This could substantially boost trials requiring large samples or slowly recruiting trials. Yet, the ability to assess outcomes without having to measure or collect them could be the most appealing resource-sparing advantage of EHRs in RCTs. Even when funds are not at issue, just the decrease in logistical difficulties themselves, particularly in large RCTs, could be worth extracting routinely collected EHR data. Thus, the EHR may have an important role in the potential for implementing large and pragmatic trials.^{2,3} This offers entirely new perspectives on evaluating health care interventions that favour the development of learning health care systems.⁷

However, the cost associated with implementing the EHR/electronic medical record infrastructure may be substantial.⁹ Although one could argue that using EHRs for research purposes might lead to more affordable trials, to our knowledge, there is no systematic overview of empirical cost estimates per individual trial participant in EHR-supported RCTs. We conducted a systematic descriptive survey of the use of EHRs in RCTs to determine how EHRs are implemented in clinical research settings and to describe specifically how this technology is used to support recruitment and outcome assessment. We aimed to determine the frequency of use of EHRs and describe possible applications of EHR technology in current practice, focusing on trials that were supported by the EHR rather than evaluating the EHR itself. We also aimed to determine the cost of using EHRs for RCTs.

Methods

We performed a descriptive study assessing the current use of EHR technology in RCTs. We included any RCT in humans published in English since January 2000 that addressed any health-related topic and that used EHRs for any purpose, including participant recruitment, intervention delivery or outcome assessment.¹⁰ Focusing on modern technology, we did not include older trials. There were no other eligibility criteria.

Definitions of the EHR and related data vary.^{10–12} Our working definitions are shown in Appendix 1. Briefly, we considered EHRs an archive of health-related data in digital form, collected during routine clinical care for each individual patient, stored and exchanged securely and accessible by multiple authorized users in a network of care providers.¹¹ The EHR infrastructure used in eligible RCTs must have already existed, and data must have been obtained through a query of the EHR database (i.e., no data fabricated specifically for the experiment would be considered routinely collected, for

example, when the trial was about the novel implementation of an EHR v. no such implementation). There is no protocol published for this descriptive study.

Literature search

We searched PubMed (last search on Sept. 13, 2017) for English-language articles published since Jan. 1, 2000 using keywords including “electronic health record,” “electronic medical record,” “health information exchange,” “patient health record” and “e-health” with an established RCT filter¹³ (Appendix 2, available at www.cmajopen.ca/content/7/1/E23/suppl/DC1). Our search integrated the search strategy for EHRs provided by the United States National Library of Medicine¹⁴ and was developed with the support of an information specialist (H.E.). Two reviewers (K.A.M. and H.E. or A.L.) screened titles and abstracts. We obtained any article deemed pertinent by at least 1 reviewer as full text. One reviewer (K.A.M.) evaluated full texts and determined eligibility, and another reviewer (L.G.H.) confirmed all exclusions.

Data extraction

We classified eligible RCTs based on the way in which the EHRs were used for 1) patient recruitment in any form, 2) outcome assessment in any form, 3) the trial intervention itself or 4) other possible purposes. For patient recruitment, we considered any effort of identifying trial participants based on certain characteristics that was done through an EHR query, as well as any random allocation of consecutive patients done through the EHR. For outcome assessment, we considered any trial in which any of the outcomes was obtained by querying or manually checking the EHR document (thus, where the outcome was routinely found within the EHR).

We then subclassified included RCTs into 1) EHR-supported trials, in which the EHR was used as research tool for conducting the trial (e.g., when patients with certain conditions are identified as enhanced recruitment strategy or adverse outcomes are queried through a hospital) and 2) EHR-evaluating trials, in which use of an EHR or an EHR modification was evaluated as part of the randomly allocated intervention (e.g., software alteration or addition, such as randomized implementation of a drug interaction alert system in a hospital's EHR ordering system). Furthermore, we extracted the RCT's research question, other study characteristics (sample size, country of origin and unit of randomization) and whether the trial included order entry systems (computerized physician order entry system or clinical decision-support system) (see Appendix 1), telehealth or personal health records.

For EHR-supported trials, we also determined the trial setting and more specific EHR uses (type of EHR and application in the trial, such as the type of alerts it would display in decision-support systems). Furthermore, we extracted whether an advanced algorithm for patient identification/recruitment or other purpose was developed. We also recorded whether the recruitment was done prospectively (e.g., by advertisement and invitation, not through the EHR), concurrently (i.e., in the point-of-care setting, through the

EHR) or retrospectively (i.e., screening a patient list, through the EHR or not), and whether routinely collected data were the only outcome source or whether a hybrid approach was used. A hybrid approach could be that 1) some outcomes were based on routinely collected data alone and other outcomes were entirely actively collected or 2) some outcomes were measured based on routinely collected data, and this measurement was supplemented by active data collection (e.g., when reported by patients outside an EHR network), or a relevant amount (more than 10% of the total routinely collected data source) was manually checked for validation. We specifically recorded the primary outcome of the trial and whether it was measured with the use of routinely collected EHR data alone, when it was measured (duration of follow-up), and any information on missing data or loss to follow-up. Furthermore, we extracted, for each trial, whether blinding and allocation concealment measures were performed. We searched the full texts for keywords, such as “placebo,” “blind,” “label” and “mask,” to identify such statements and then proceeded with extracting the statement when reported. One reviewer (K.A.M.) extracted all data. We aimed to provide a general overview on potential issues of bias in the EHR-supported studies. Two reviewers (K.A.M. and B.S.) used the Cochrane risk of bias tool,¹⁵ and a third reviewer (K.A.M., H.E., B.S. or L.G.H.) verified the assessments. Any disagreement was resolved by discussion.

Trial costs

We contacted the authors of included EHR-supported trials through a standardized email to request cost information and extracted any cost information reported in the publications. We aimed to obtain a cost estimate that would allow comparison with traditional trials. Therefore, we were not interested in costs of EHR-evaluating trials. We explained to the authors that the costs of the trial could have been divided in 3 major ways:¹⁶ 1) cost of the project/trial development and preparation (e.g., insurance, travel, infrastructure, consulting, sample size calculation, database set-up), 2) cost of enrolment, treatment and follow-up (e.g. per-patient costs, salary costs, patient reimbursement costs, material and/or drug costs) and 3) cost after last patient out (e.g., data-cleaning costs, analysis costs, publication costs). We aimed for only a raw cost estimate and accepted any information we could. We converted cost values to US dollars where applicable, based on the exchange rate on Nov. 1, 2017.¹⁷ We sent the data presented here to all trial authors for confirmation.

All costs are reported in US dollars.

Statistical analysis

We report results descriptively using proportions and medians with interquartile ranges. Since our study was exploratory, we did not use any statistical tests.

Ethics approval

As this article does not contain any personal medical information about any identifiable living people, ethics approval was not required.

Results

After 1680 titles and abstracts were screened, 394 potentially relevant articles were obtained as full texts, of which 189 were eligible (Figure 1). Of the 189 RCTs, 17 (9.0%) were supported by EHRs, and in 172 (91.0%), EHRs were used as a modality of intervention (EHR-evaluating).

Most of both EHR-supported and EHR-evaluating trials originated from North America (13 [76%] and 140 [81.4%], respectively) and were published recently (median year 2012 [interquartile range 2009–2014]) (Table 1). Three (18%) of the EHR-supported trials and 61 (35.5%) of the EHR-evaluating trials were cluster-randomized. There were no placebo-controlled trials in our sample. The majority (101 [53.4%]) did not report the level of blinding. Blinded outcome assessment was the most frequent type of blinding (35 trials [18.5%]), followed by open label (27 [14.3%]), single-blinding (19 [10.0%]) and double-blinding (7 [3.7%]).

Trials supported by electronic health records

The interventions and settings varied among the 17 EHR-supported RCTs^{18–34} (Table 2). Five trials (29%) used the EHR of a US Department of Veterans Affairs or affiliated facility. Most trials evaluated quality-improvement interventions, which often involved clinician education and feedback

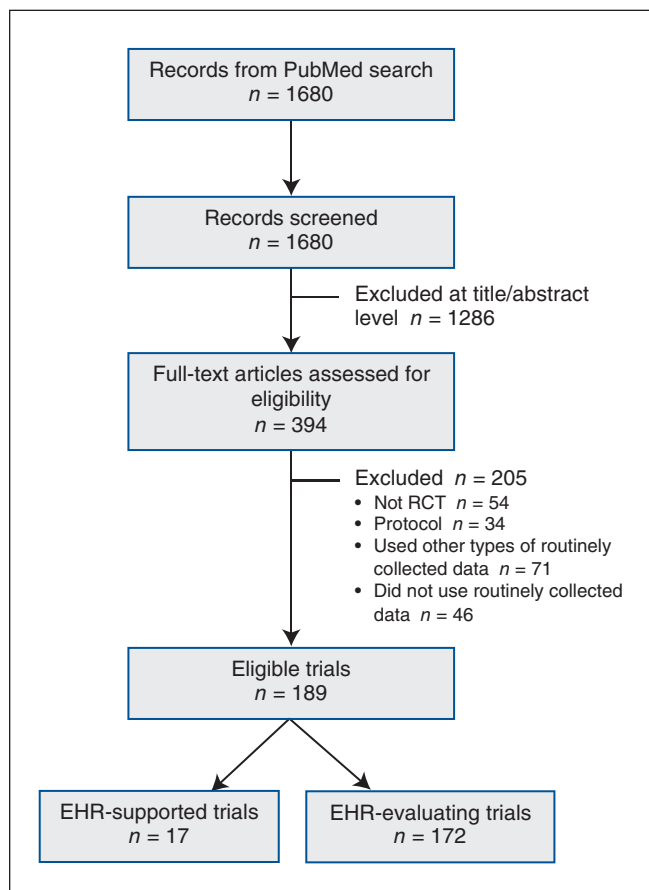


Figure 1: Flow chart showing trial selection. Note: EHR = electronic health record, RCT = randomized controlled trial.

Table 1: Characteristics of randomized controlled trials published in English between January 2000 and Sept. 13, 2017 that used electronic health records

Characteristic	No. (%) of trials*		
	Overall n = 189	EHR-evaluating n = 172	EHR-supported n = 17
EHR for intervention	172 (91.0)	172 (100.0)	–
Computerized physician order entry system or clinical decision-support system	128 (67.7)	128 (74.4)	–
Telehealth	14 (7.4)	14 (8.1)	–
Personal health record	26 (13.8)	26 (15.1)	–
Electronic patient-reported outcomes	4 (2.1)	4 (2.3)	–
EHR for outcome measurement	158 (83.6)	143 (83.1)	15 (88)
EHR for patient recruitment	105 (55.6)	91 (52.9)	14 (82)
Country/region			
North America	153 (81.0)	140 (81.4)	13 (76)
United Kingdom	9 (4.8)	7 (4.1)	2 (12)
Continental Europe	15 (7.9)	14 (8.1)	1 (6)
Other†	12 (6.3)	11 (6.4)	1 (6)
Cluster RCT	64 (33.9)	61 (35.5)	3 (18)
Unit of randomization			
Clinician	49 (25.9)	46 (26.7)	3 (18)
Patient	76 (40.2)	65 (37.8)	11 (65)
Pharmacy	1 (0.5)	1 (0.6)	0 (0)
Practice/clinic	54 (28.6)	51 (29.6)	3 (18)
Unit/floor	9 (4.8)	9 (5.2)	0 (0)
Publication year, median (IQR)	2012 (2009–2014)	2012 (2009–2014)	2013 (2010–2013)
Sample size, median (IQR)			
Total	89 (24–732)	80 (22–513)	732 (73–2513)
Cluster RCTs excluded	239 (57–1187)	254 (60–1187)	900 (111–3075)
Cluster RCTs only	24 (12–47)	24 (12–52)	18 (12–24)
Blinding			
Open label	27 (14.3)	23 (13.4)	4 (24)
Single-blinded	19 (10.0)	18 (10.5)	1 (6)
Double-blinded	7 (3.7)	6 (3.5)	1 (6)
Outcome assessment blinding	35 (18.5)	30 (17.4)	5 (29)
Not reported	101 (53.4)	95 (55.2)	6 (35)

Note: EHR = electronic health record, IQR = interquartile range, RCT = randomized controlled trial.
 *Except where noted otherwise.
 †Includes China, Japan, Taiwan, Iran, India, Pakistan, Lebanon, Australia and Kenya.

initiatives (8 [47%]), screening programs (4 [24%]), and collaborative care and disease management interventions integrated into primary care settings (3 [18%]). Almost half of the studies (8 [47%]) took place in primary care clinics, 5 (29%) were conducted in health care networks, and 3 (18%) took place in hospitals. One trial¹⁸ (6%) was performed entirely within a pharmacy electronic medical record.

Supported outcome measurement

Fifteen trials (88%) measured outcomes using the EHR (Table 1). The EHR-assessed outcomes were typically screening uptake (e.g., women seeking a Papanicolaou test after receiving an automated call from the EHR prompting cervical cancer screening) (6 trials [40%]), clinical outcomes (4 [27%]), drug adherence (2 [13%]) or guideline-concordant

Table 2 (part 1 of 2): Characteristics of randomized controlled trials supported by electronic health records

Investigator	Country; sample size	EHR use for recruitment; study type	EHR use for outcome assessment; extent of use of routinely collected data	Patient population/ indication	Intervention and control*	Primary outcome; length of follow-up; amount of missing data	Setting
Bereznicki et al., ¹⁹ 2008	Australia; 1551 patients	Yes; retrospective	Yes; EHR alone	Uncontrolled asthma	Intervention: contact by community pharmacist plus educational material and referral to general practitioner for asthma management	Ratio of dispensed preventer and reliever medication; 6 mo; NR	Community pharmacy network
Corson et al., ¹⁹ 2011	US; 42 caregivers (randomly allocated), 365 patients	No; prospective	Yes; hybrid, primary outcome EHR alone	Musculoskeletal pain	Intervention: patient and clinician education, symptom monitoring and feedback to clinicians	Guideline-concordant care; 12 mo; NR	Primary care clinics associated with Department of Veterans Affairs medical centre and urban hospital
de Jong et al., ²⁰ 2013	Netherlands; 73 general practitioner trainees (randomly allocated)†	Yes; retrospective	Yes; hybrid, primary outcome EHR alone	Skin and psychosocial conditions	Steering patient mix of general practitioner trainees	Trainee exposure to specific field, and knowledge and self-efficacy; 6 mo; 5%–10%	Practice network with general practitioner training program
Fu et al., ²¹ 2014	US; 6400 patients	Yes; retrospective	Yes, hybrid; primary outcome active data collection alone	Current smokers	Proactive outreach plus choice of smoking cessation services	6-mo prolonged smoking abstinence at 1 yr; 12 mo; 48.3% (but 0% for EHR outcome)	Department of Veterans Affairs medical centre
Galbreath et al., ²² 2004	US; 1069 patients	Yes; retrospective	Yes, hybrid; primary outcome active data collection alone	Symptomatic congestive heart failure	Congestive heart failure management program (plus at-home scale)	All-cause mortality and use of health care services; NR (time to event); NR	Various health care networks‡ and Medicare/Medicaid participants
Gerber et al., ²³ 2013	US; 18 practices, 170 caregivers (randomly allocated), 185 212 patients	Yes; retrospective	Yes, hybrid; primary outcome active data collection alone	Clinical practice groups with primary care pediatricians (children with acute respiratory tract infections)	Antibiotic stewardship program	Change in broad-spectrum antibiotics prescribed for bacterial infections or in antibiotic prescribed for viral infections; 12 mo; 5% of caregivers	Pediatric primary care network
Green et al., ²⁴ 2013	US; 4675 patients	Yes; retrospective	Yes; EHR alone	Prevention of colorectal cancer	Automated interventions v. assisted care v. navigated care v. usual care	Receiving any colorectal cancer test and being current for colorectal cancer testing in years 1 and 2; 24 mo; 0.2%	Primary care practice network
Hoffman et al., ²⁵ 2010	US; 404 patients	Yes; retrospective	No; active data collection alone	Prevention of colorectal cancer	Fecal immunochemical test (v. guaiac-based occult blood test)	Screening adherence; 3 mo; NR	Department of Veterans Affairs network (primary care clinics and laboratory)
Israel et al., ²⁶ 2013	US; 732 patients	Yes, retrospective	No; active data collection alone	Adult inpatients with at least 1 of several cardiovascular disease diagnoses in EHR	Minimal intervention (medication reconciliation), enhanced intervention (minimal intervention plus pharmacist) or usual care	Rate of underuse of cardiovascular drugs; 3 mo; NR	University hospital (orthopedic, internal medicine, family medicine and cardiology wards)

Table 2 (part 2 of 2): Characteristics of randomized controlled trials supported by electronic health records

Investigator	Country; sample size	EHR use for recruitment, type	EHR use for outcome assessment; extent of routinely collected data use	Patient population/ indication	Intervention and control*	Primary outcome; follow-up; missing data	Setting
McCarren et al., ²⁷ 2013	US; 12 practices (randomly allocated), 220 patients	Yes; retrospective	Yes; EHR alone	Heart failure and guideline-nonconcordant β-blocker prescription	Information to pharmacy about prescription nonconcordance	Guideline-concordant prescriptions; 6 mo; 0%	Veterans Health Administration facilities and pharmacies
Phillips et al., ²⁹ 2011	US; 3895 patients	Yes; retrospective	Yes; EHR alone	Prevention of breast cancer	Telephone calls and reminder letters from patient navigators	Adherence to biennial mammography; 9 mo; NR	Hospital-based internal medicine practices
Piazza et al., ³⁰ 2013	US; 2513 patients	Yes; retrospective	Yes; EHR alone	Medical service inpatients at risk for venous thromboembolism with planned discharge within 48 h	Alert for physician	Symptomatic deep vein thrombosis or pulmonary embolism; 3 mo; < 0.1%	Inpatient medical unit
Qureshi et al., ³¹ 2012	UK; 24 caregivers (randomly allocated), 748 patients	No; prospective	Yes; hybrid, primary outcome EHR alone	Adult primary care patients with no previously diagnosed cardiovascular risk	Family history questionnaire (in addition to Framingham risk score)	Proportion of identified participants with high cardiovascular risk scores; NA; 1.7%	Family practices in research network
Skinner et al., ³² 2015	US; 1032 patients	Yes; retrospective	Yes; EHR alone	Prevention of colorectal cancer	Tablet-based Cancer Risk Intake System assessment before appointment and control group	Received risk-appropriate colorectal cancer testing and any type of colorectal cancer testing; 12 mo; 0%	Family practices affiliated with university medical centre
Stewart et al., ²⁸ 2014	US; 235 patients	Yes; retrospective	Yes; hybrid, primary outcome EHR alone	Dysthymia or major depressive disorder	Collaborative care program with psychotherapy and antidepressant drugs	Cardiovascular events; 96 mo; 0%	Academic group practice
Vestbo et al., ³³ 2016	UK; 2802 patients	No; prospective	Yes; hybrid, primary outcome active data collection alone	COPD and regular maintenance inhaler therapy	Fluticasone furoate + vilanterol (100 µg/25 µg) once a day via inhalation v. usual care	Moderate or severe COPD exacerbation; 12 mo; 24.8%	Health care network in and around Salford, hospitals, general practitioners, pharmacies
Wolf et al., ³⁴ 2005	US; 113 health care providers (randomly allocated), 1978 patients	Yes; retrospective	Yes; EHR alone	Prevention of colorectal cancer	Education session plus performance feedback	Completion of colorectal cancer screening; NA; NR	Department of Veterans Affairs primary care clinics

Note: COPD = chronic obstructive pulmonary disease, EHR = electronic health record, NA = not applicable, NR = not reported.

*All comparisons are "usual care" unless otherwise specified.

†Number of patients not reported.

‡University of Texas Health Science Center at San Antonio, in partnership with Wilford Hall Medical Center, Brooke Army Medical Center, South Texas Veterans Health Care System, TRICARE Region 6 and University Health System.

care measures (2 [13%]). In 7 (47%) of the 15 trials, the routinely collected data source was the only source of outcome data in the entire trial. In the remaining 8 trials (53%), a hybrid approach was applied, with some outcome data being collected actively. In 4^{19,20,28,31} of these 8 cases, the primary outcome was fully extracted from an EHR but additional outcomes were collected actively, and in 3 cases,^{21,22,33} the pri-

mary outcome was collected actively but additional outcomes were EHR-based. In 1 case,²³ the primary outcome was collected through the EHR but was verified with actively collected data. Twelve (80%) of the 15 trials relied on the EHR for primary outcome assessment.

The median trial duration was 10 months (interquartile range 5–12 mo); in 10 (59%) of the 17 trials, the number of

missing data or patients lost to follow-up was reported, but none reported on the quality of the data.

Supported recruitment

Fourteen trials (82%) used the EHR as a tool for patient recruitment (Table 1). A prospective approach was reported in 1 trial,³⁰ and in the remaining 13 trials, the EHR was used retrospectively (i.e., manual check or simple retrospective query of eligible patients via EHR). In addition, 1 trial¹⁸ used a complex querying system (another trial²⁷ appeared to, but this was not specifically reported). The remaining 3 trials (18%) used a (traditional) prospective recruitment approach without the use of EHRs.

Costs

We contacted 13 of the 17 corresponding authors of the EHR-supported trials. Emails were undeliverable to 3 addresses, for which we were also unable to find an alternative contact online, and we were not able to reach the authors. We obtained information on trial costs for 4^{18,24,27,30} of the 17 trials and intervention cost data for 1 trial³⁴ (response rate 24%).

Cost information came from 1 Australian trial¹⁸ and 4 US trials^{24,27,30,34} (2 within the Department of Veterans Affairs network^{27,34}). The costs varied from \$67 750 to \$5 026 000 (median \$86 753) for total trial costs and from \$44 to \$2000 (median \$315) for per-patient costs (Table 3). Overall trial costs were derived from funding budgets in 3 cases,^{18,2,27} and 1 author stated that the overall costs were \$2000 per patient.³⁰ In the trial in which the EHR database was leveraged through automated data extraction,¹⁸ the per-patient cost was \$44. In the 2 trials in which the extraction of study data from the

EHR source was done manually,^{24,30} the per-patient cost was \$560 and \$2000, respectively. We have no information in this regard for 1 trial.²⁷ For the trial that presented only the costs of the intervention (extracting data from the EHR to give feedback to health care providers),³⁴ a cost of \$44 per patient was reported when the data were extracted manually, and a sensitivity analysis indicated that this cost could decrease to \$9 if the data were extracted automatically.

Risk of bias

Three trials had no indication for high risk of bias in any of the domains assessed^{18,25,26} (Table 4). There were no indications for high risk of bias related to randomization or allocation concealment in any of the trials. Most trials were open-label or assessed an intervention that was not disguisable from the participants/providers, which may indicate a high risk of bias. Relevant to EHR trials, the risk for attrition bias was generally low (missing outcome data for not more than 10% of patients), and in 4 trials, all data were reported for all patients.^{21,27,28,32}

Trials using electronic health records for intervention

Among the 172 EHR-evaluating trials (references in Appendix 3, available at www.cmajopen.ca/content/7/1/E23/suppl/DC1), the investigators measured outcomes using the EHR in 143 (83.1%), and the EHR was used as a tool for patient recruitment in 91 (52.9%) (Table 1). Computerized decision-support systems such as a physician order entry system or a clinical decision-support system were evaluated in 128 (74.4%). Personal health records were evaluated in 26 (15.1%).

Table 3: Costs of randomized controlled trials supported by electronic health records*

Type of data extraction from EHR source; investigator	EHR source preexisting	Intervention integrated during routine care (no additional staff needed)	Total trial cost, US\$	No. of patients†	Cost per patient, US\$
Automatic					
Bereznicki et al., ¹⁸ 2008	Yes	Yes	67 750††	1551	44
Manual					
Green et al., ²⁴ 2013	Yes	No	2 800 000§	5000	560
Piazza et al., ³⁰ 2013	Unclear	No	5 026 000¶	2513	2000
Wolf et al., ³⁴ 2005	Yes	Yes	86 753**	1978	44
Unclear whether automatic or manual					
McCarren et al., ²⁷ 2013	Yes	Yes	69 300††	220	315
Note: EHR = electronic health record. *In all trials, the EHR was used for recruitment (retrospective) and outcome assessment (all with EHR alone). †As reported by authors for these costs data. ‡Total received funding, including \$42 157 for staff costs for the duration of the project, \$6132 for a consultant programmer (for software development), \$15 330 for pharmacy payments and \$6132 for nonsalary costs such as printing, postage and travel. §Total received funding. ¶Study costs were \$2000 per patient and included costs of the trial start-up and close out. **Total cost of the colorectal cancer screening promotional effort (intervention only). ††Total received funding. "Most of the [working] time was donated."					

Table 4: Risk of bias assessment with the Cochrane risk of bias tool¹⁵

Investigator	Random sequence generation	Allocation concealment	Blinding of participants and personnel	Blinding of outcome assessment	Incomplete outcome data (> 10%)	Selective outcome reporting*
Bereznicki et al. ¹⁸	?	–	–	–	?	NA
Corson et al. ¹⁹	?	?	+	–	?	NA
de Jong et al. ²⁰	–	?	+	?	–	NA
Fu et al. ²¹	?	?	+	–	+†	NA
Galbreath et al. ²²	?	?	+	+	?	NA
Gerber et al. ²³	?	–	+	?	–	NA
Green et al. ²⁴	–	–	+	–	–	NA
Hoffman et al. ²⁵	–	?	?	?	?	NA
Israel et al. ²⁶	–	?	–	–	?	NA
McCarren et al. ²⁷	–	?	+	+	–	NA
Phillips et al. ²⁹	?	?	+	?	?	NA
Piazza et al. ³⁰	–	–	+	?	–	NA
Qureshi et al. ³¹	–	–	+	?	–	NA
Skinner et al. ³²	?	?	+	?	–	NA
Stewart et al. ²⁸	–	–	+	–	–	NA
Vestbo et al. ³³	–	–	+	?	+	NA
Wolf et al. ³⁴	?	?	+	?	?	NA

Note: EHR = electronic health record, NA = nonapplicable, – = low risk of bias, ? = unclear risk of bias, + = high risk of bias.
 *Since only the publication identified in our literature search was assessed (there was no systematic protocol search and no searching for further manuscripts related to that study), we did not consider this item.
 †The completeness of outcome data based on the electronic health record (secondary study outcome) was perfect (0% missing data).

Telemonitoring-tethered devices measuring vital signs that were connected to the EHR were evaluated in 14 trials (8.1%), and electronic patient-reported outcomes were evaluated in 4 (2.3%) (Table 1).

Interpretation

In most of the identified trials in which EHRs were used, EHR technology itself was explored. However, we identified 17 trials that investigated an EHR-unrelated intervention and were supported by the use of EHRs for patient recruitment or outcome assessment. Most trials were published recently, indicating a rapid development in this field.

The potential of registry-based trials for comparative effectiveness research and the current state of using registries for RCTs, in particular for outcome ascertainment, has been reviewed recently.^{8,35} Interestingly, although the settings and implementation were similar to those identified in our sample, registry trials are most frequently performed in Scandinavian countries³⁵ and EHR trials predominantly in North America. In addition, primary outcome data in registry trials are often collected with the use of routine data (82%), similarly to EHR trials (80%), which indicates confidence in the reliability of these data.³⁵ Information about data quality and validity was rarely reported for registry-based trials (11%)³⁵ and was not reported in any of the EHR-supported RCTs in our sample,

which suggests similar reporting problems as in observational research based on routinely collected data.³⁶ This may be expected given the current lack of a standardized reporting guideline for RCTs in which routinely collected data are used but also highlights a substantial transparency problem.

In most (84%) of the trials in our sample, outcomes were measured with the use of EHRs, including many of the most patient-relevant clinical outcomes, from unscheduled hospital admission to death. But there were also less pragmatic and more exploratory, mechanistic^{37,38} outcomes that help to understand pathophysiological processes: for example, in 1 study, EHR-extracted lipid levels were used during a trial of a lipid-lowering agent.³⁹ We also identified a trial, the Salford lung study,³³ that used routinely collected data in a prelicensing setting in the context of drug approval.

The identified EHR-supported trials were heterogeneous with regard to their targeted populations and outcomes measured, with a few exceptions. For example, over a third of this subsample were Veterans Affairs trials, in which the EHR was used for outcome and patient identification in all cases. This is likely due to the fact the Veterans Affairs has a long-established EHR system, and its widespread network allows for ease in designing and implementing these types of trials.

Another interesting finding that relates to the EHR-evaluating trials in our sample is the high proportion (about one-third) of trials in which cluster randomization was used. This indicates that

EHR-based trials mostly evaluate interventions not at the patient level but, rather, more at a system level, as when aiming to redirect physician behaviour. This introduces the risk of contamination between the units of randomization (e.g., physicians) and thus requires a cluster design to be implemented.

Other than affordability, the great theoretical value of integrating the EHR into clinical trials lies in its potential for patient recruitment. For example, D'Avolio and colleagues⁴⁰ reported on a Veterans Affairs pilot study that, like those identified in our sample, showed how convenient it can be to identify patients based on specific characteristics (the EHR database is "scanned," and a list of possibly eligible patients results) and even to recruit them by sending an automatic electronic message to their clinician. Even with a lower response rate, when the contacted patients are in the order of thousands, this could lead to greater recruitment capacity, which could be of substantial value, particularly in RCTs in which difficult recruitment is suspected during planning. We found that, in almost half of the EHR-supported trials that used EHRs for recruitment, the investigators made use of more sophisticated techniques such as the proposed mechanisms of data mining. Although in some trials, patients were recruited by screening the EHR without specifying the use of a particular algorithm in addition, most EHRs will require some programming to identify specific traits in the system that go beyond the basic EHR abilities (e.g., typing an International Statistical Classification of Diseases and Related Health Problems, 10th revision, Clinical Modification code for diabetes in a search window to obtain a list of patients, which can be done manually). More advanced EHR add-ons, which can screen for multiple variables at multiple levels simultaneously and continuously (i.e., screening the system every 2 hours or instantly during care for the entire time of the trial) require planning and validation. An example of such an EHR screening tool is one developed and used in the 2008 trial by Bereznicki and colleagues,¹⁸ in which this data-mining tool scrutinized the pharmacy electronic medical record based on a specified protocol (history of asthma medication's being dispensed more frequently than guideline customs) to flag patients with poorly controlled asthma. These patients were then contacted, received educational material for self-management and were prompted to contact their care providers. This shows how using the EHR for patient identification and recruitment can be done efficiently yet requires substantial planning and software development. We provide a general framework with the various potential applications and challenges of using routinely collected data in different trial conduct phases elsewhere.^{3,6}

The author-reported costs could support the assumption that using routinely collected data for RCTs may promote cost reduction as long as the outcome data source is already established and is not a financial responsibility of the research endeavour. In the 3 trials in which the EHR infrastructure was well established and was merely redirected for use in the trials,^{18,27,34} the cost per patient (median \$44) was much lower than often-reported costs in traditional trials.⁴¹ The costs of the 2 trials in which the infrastructure was less integrated (such as actively screening the EHR for assessing the clinical outcome),

a median of \$1280 per patient, were more similar to those of traditional RCTs.^{24,30} A recently published overview of registry trials showed similar trial cost patterns (i.e., a reduction of costs when the outcome data did not require manual collection but, rather, the registry infrastructure was leveraged).⁸

Limitations

Some limitations of our study merit attention. First, we did not aim for a complete sample of all published EHR-based trials, and we searched PubMed only; rather, we aimed for a systematic, comprehensible and reproducible survey of the current literature. We used a highly sensitive search algorithm and implemented specific EHR search filters provided by the US National Library of Medicine. Nonetheless, we assume that we overlooked several pertinent publications that did not indicate in their keywords, title or abstract the application of EHRs. This may have engendered overrepresentation of EHRs used for interventions in our sample, and the observed disproportion of EHR-evaluating and EHR-supported trials needs to be interpreted with caution.

Second, searching for English-language articles indexed in PubMed alone may have created regional bias, with potential overrepresentation of Anglo-American studies. This could explain the high proportion of studies from the US. Nonetheless, substantial legislative and financial efforts have been made in North America to encourage the acquisition and employment of EHR technology, which may more likely be the reason for this critical imbalance.

Third, the trials were highly diverse, showing the various fields of EHR application, but we would need more data to further evaluate individual details and to explore, for example, the ethical constraints associated with no-consent point-of-care trials.^{42,43}

Fourth, only 1 reviewer assessed the eligibility of the full texts and completed several parts of the data extraction, which may have introduced error in the selection of the trials. Nonetheless, we feel that the identified trials provide an overview of the mode of use of the EHR in RCTs.

Fifth, we did not test any hypothesis regarding the effect of using the EHR in trials, nor did we assess the impact of using the EHR on outcome ascertainment. Although we extracted a few characteristics that can point to the methodological quality of the studies, including an evaluation of major domains of risk of bias, we did not evaluate the treatment effects reported in the trials but merely offered a description of their use.

Finally, we obtained only a few rough cost estimates without details, which did not allow us to deduce any cost patterns; however, it provided first estimates to shed some light in this area.

Conclusion

Electronic health records are a novel and valuable addition to clinical research. There are numerous examples of how the EHR was implemented successfully in clinical research settings, supporting recruitment and outcome measurement in randomized trials. Electronic health records may be associated with lower research costs, allowing the conduct of more or larger RCTs. Altogether, these are promising developments toward more randomized real-world evidence.

References

1. Bothwell LE, Greene JA, Podolsky SH, et al. Assessing the gold standard — lessons from the history of RCTs. *N Engl J Med* 2016;374:2175-81.
2. Zuidgeest MGP, Goetz I, Groenwold RHH, et al.; GetReal Work Package 3. Series: Pragmatic trials and real world evidence: Paper 1. Introduction. *J Clin Epidemiol* 2017;88:7-13.
3. Mc Cord KA, Al-Shahi Salman R, Treweek S, et al. Routinely collected data for randomized trials: promises, barriers, and implications. *Trials* 2018;19:29.
4. Ramsberg J, Neovius M. Register or electronic health records enriched randomized pragmatic trials: The future of clinical effectiveness and cost-effectiveness trials? *Nordic J Health Econ* 2015;3:1-15.
5. PCORI-funded Pragmatic Clinical Studies projects. Washington: Patient-Centered Outcomes Research Institute; 2016 [updated 2017 Sept. 26]. Available: www.pcori.org/research-results/pragmatic-clinical-studies/pcori-funded-pragmatic-clinical-studies-projects (accessed 2017 Oct. 10).
6. Mc Cord KA, Hemkens LG. Using electronic health records for clinical trials: Where do we stand and where can we go? *CMAJ* 2019;191:E128-33.
7. Friedman CP, Wong AK, Blumenthal D. Achieving a nationwide learning health system. *Sci Transl Med* 2010;2:57cm29.
8. Li G, Sajobi TT, Menon BK, et al.; 2016 Symposium on Registry-Based Randomized Controlled Trials in Calgary. Registry-based randomized controlled trials — What are the advantages, challenges, and areas for future research? *J Clin Epidemiol* 2016;80:16-24.
9. Menachemi N, Collum TH. Benefits and drawbacks of electronic health record systems. *Risk Manag Healthc Policy* 2011;4:47-55.
10. What is an electronic health record (EHR)? Washington: Office of the National Coordinator for Health Information Technology; [reviewed 2018 Mar. 21]. Available: www.healthit.gov/faq/what-electronic-health-record-ehr (accessed 2018 Dec. 10).
11. *Health informatics — electronic health record — definition, scope and context*. ISO/TR 20514:2005. Geneva: International Organization for Standardization; 2005.
12. Electronic health records. Baltimore: Centers for Medicare & Medicaid Services; [modified 2012 Mar. 26]. Available: www.cms.gov/Medicare/E-Health/EHealthRecords/index.html?redirect=/EhealthRecords/ (accessed 2017 Nov. 10).
13. Higgins JPT, Green S, editors. Box 6.4.a: Cochrane highly sensitive search strategy for identifying randomized trials in MEDLINE: sensitivity-maximizing version (2008 revision); PubMed format. In: *Cochrane handbook for systematic reviews of interventions*. Version 5.1.0. Oxford (UK): Cochrane Collaboration; 2011. Available: <http://handbook-5-1.cochrane.org> (accessed 2017 Jan. 1).
14. 2016 MeSH highlights: questions and answers. Bethesda (MD): U.S. National Library of Medicine; 2016 [updated 2018 Apr. 27]. Available: www.nlm.nih.gov/bsd/disted/clinics/mesh_2016_qa.html (accessed 2017 Nov. 10).
15. Higgins JPT, Green S, editors. The Cochrane Collaboration's tool for assessing risk of bias. In: *Cochrane handbook for systematic reviews of interventions*. Version 5.1.0. Oxford (UK): Cochrane Collaboration; 2011. Available: <http://handbook-5-1.cochrane.org> (accessed 2019 Jan. 8).
16. Speich B, von Niederhäusern B, Blum CA, et al.; MAKing Randomized Trials Affordable (MARTA) Group. Retrospective assessment of resource use and costs in two investigator-initiated randomized trials exemplified a comprehensive cost item list. *J Clin Epidemiol* 2018;96:73-83.
17. Realtimecourse | Aktien | Börsenkurse | Börse [exchange rates]. Available: www.finanzen.ch/ (accessed 2017 Nov. 1).
18. Bereznicki BJ, Peterson GM, Jackson SL, et al. Data-mining of medication records to improve asthma management. *Med J Aust* 2008;189:21-5.
19. Corson K, Doak MN, Denneson L, et al. Primary care clinician adherence to guidelines for the management of chronic musculoskeletal pain: results from the study of the effectiveness of a collaborative approach to pain. *Pain Med* 2011;12:1490-501.
20. de Jong J, Visser MR, Wieringa-de Waard M. Steering the patient mix of GP trainees: results of a randomized controlled intervention. *Med Teach* 2013;35:101-8.
21. Fu SS, van Ryn M, Sherman SE, et al. Proactive tobacco treatment and population-level cessation: a pragmatic randomized clinical trial. *JAMA Intern Med* 2014;174:671-7.
22. Galbreath AD, Krasuski RA, Smith B, et al. Long-term healthcare and cost outcomes of disease management in a large, randomized, community-based population with heart failure. *Circulation* 2004;110:3518-26.
23. Gerber JS, Prasad PA, Fiks AG, et al. Effect of an outpatient antimicrobial stewardship intervention on broad-spectrum antibiotic prescribing by primary care pediatricians: a randomized trial. *JAMA* 2013;309:2345-52.
24. Green BB, Wang CY, Anderson ML, et al. An automated intervention with stepped increases in support to increase uptake of colorectal cancer screening: a randomized trial. *Ann Intern Med* 2013;158:301-11.
25. Hoffman RM, Steel S, Yee EF, et al. Colorectal cancer screening adherence is higher with fecal immunochemical tests than guaiac-based fecal occult blood tests: a randomized, controlled trial. *Prev Med* 2010;50:297-9.
26. Israel EN, Farley TM, Farris KB, et al. Underutilization of cardiovascular medications: effect of a continuity-of-care program. *Am J Health Syst Pharm* 2013;70:1592-600.
27. McCarren M, Furmaga E, Jackevicius CA, et al. Improvement of guideline β -blocker prescribing in heart failure: a cluster-randomized pragmatic trial of a pharmacy intervention. *J Card Fail* 2013;19:525-32.
28. Stewart JC, Perkins AJ, Callahan CM. Effect of collaborative care for depression on risk of cardiovascular events: data from the IMPACT randomized controlled trial. *Psychosom Med* 2014;76:29-37.
29. Phillips CE, Rothstein JD, Beaver K, et al. Patient navigation to increase mammography screening among inner city women. *J Gen Intern Med* 2011;26:123-9.
30. Piazza G, Anderson FA, Ortel TL, et al. Randomized trial of physician alerts for thromboprophylaxis after discharge. *Am J Med* 2013;126:435-42.
31. Qureshi N, Armstrong S, Dhiman P, et al.; ADDFAM (Added Value of Family History in CVD Risk Assessment) Study Group. Effect of adding systematic family history enquiry to cardiovascular disease risk assessment in primary care: a matched-pair, cluster randomized trial. *Ann Intern Med* 2012;156:253-62.
32. Skinner CS, Halm EA, Bishop WP, et al. Impact of risk assessment and tailored versus nontailored risk information on colorectal cancer testing in primary care: a randomized controlled trial. *Cancer Epidemiol Biomarkers Prev* 2015;24:1523-30.
33. Vestbo J, Leather D, Diar Bakerly N, et al. Effectiveness of fluticasone furoate-inhaler for COPD in clinical practice. *N Engl J Med* 2016;375:1253-60.
34. Wolf MS, Fitzner KA, Powell EF, et al. Costs and cost effectiveness of a health care provider-directed intervention to promote colorectal cancer screening among veterans. *J Clin Oncol* 2005;23:8877-83.
35. Mathes T, Buehn S, Prengel P, et al. Registry-based randomized controlled trials merged the strength of randomized controlled trials and observational studies and give rise to more pragmatic trials. *J Clin Epidemiol* 2018;93:120-7.
36. Hemkens LG, Benchimol EL, Langan SM, et al. The reporting of studies using routinely collected health data was often insufficient. *J Clin Epidemiol* 2016;79:104-11.
37. Schwartz D, Lellouch J. Exploratory and pragmatic attitudes in therapeutical trials. *J Chronic Dis* 1967;20:637-48.
38. Karanicolos PJ, Montori VM, Devereaux PJ, et al. A new “mechanistic-practical” framework for designing and interpreting randomized trials. *J Clin Epidemiol* 2009;62:479-84.
39. Lester WT, Grant RW, Barnett GO, et al. Randomized controlled trial of an informatics-based intervention to increase statin prescription for secondary prevention of coronary disease. *J Gen Intern Med* 2006;21:22-9.
40. D'Avolio L, Ferguson R, Goryachev S, et al. Implementation of the Department of Veterans Affairs' first point-of-care clinical trial. *J Am Med Inform Assoc* 2012;19:e170-6.
41. Speich B, von Niederhäusern B, Schur N, et al.; MAKing Randomized Trials Affordable (MARTA) Group. Systematic review on costs and resource use of randomized clinical trials shows a lack of transparent and comprehensive data. *J Clin Epidemiol* 2018;96:1-11.
42. Edwards SJ, Lilford RJ, Braunholtz DA, et al. Ethical issues in the design and conduct of randomised controlled trials. *Health Technol Assess* 1998;2:i-vi, 1-132.
43. Baum M. Do we need informed consent? *Lancet* 1986;2:911-2.

Affiliations: Basel Institute for Clinical Epidemiology and Biostatistics (Mc Cord, Ewald, Ladanie, Briel, Speich, Bucher, Hemkens), Department of Clinical Research, University Hospital Basel, University of Basel; University Medical Library (Ewald), University of Basel; Swiss Tropical and Public Health Institute (Ladanie), University of Basel, Basel, Switzerland; Department of Health Research Methods, Evidence, and Impact (Briel), McMaster University, Hamilton, Ont.

Contributors: Lars Hemkens conceived and designed the study. Kimberly Mc Cord, Hannah Ewald and Aviv Ladanie screened titles, abstracts and full-text publications. Kimberly Mc Cord extracted the data, and Kimberly Mc Cord and Lars Hemkens analyzed the data. Kimberly Mc Cord and Lars Hemkens drafted the manuscript. All of the authors interpreted the data, critically revised the manuscript for important intellectual content, gave final approval of the version to be published and agreed to be accountable for all aspects of the work.

Funding: This work was supported by the Stiftung Institut für klinische Epidemiologie.

Disclaimer: The funder had no role in the design and conduct of the study; collection, management, analysis and interpretation of the data; and preparation, review or approval of the manuscript or its submission for publication.

Supplemental information: For reviewer comments and the original submission of this manuscript, please see www.cmajopen.ca/content/7/1/E23/suppl/DC1.