Leveraging Real-World Data for the Selection of Relevant Eligibility Criteria for the Implementation of Electronic Recruitment Support in Clinical Trials

Georg Melzer[1]   Tim Maiwald[2]   Hans-Ulrich Prokosch[1]   Thomas Ganslandt[1,3]

[1] Chair of Medical Informatics, Friedrich-Alexander University Erlangen-Nürnberg, Erlangen, Germany
[2] Institute for Electronics Engineering, Department Electrical Engineering, Friedrich-Alexander University Erlangen-Nürnberg, Erlangen, Germany
[3] Heinrich-Lanz-Center for Digital Health, Department of Biomedical Informatics, Mannheim University Medicine, Ruprecht-Karls-University Heidelberg, Mannheim, Germany

Address for correspondence Georg Melzer, Chair of Medical Informatics, Friedrich-Alexander University Erlangen-Nürnberg, Wetterkreuz 15, 91058 Erlangen-Tennenlohe, Germany (e-mail: georg.melzer@fau.de).

Abstract

Background   Even though clinical trials are indispensable for medical research, they are frequently impaired by delayed or incomplete patient recruitment, resulting in cost overruns or aborted studies. Study protocols based on real-world data with precisely expressed eligibility criteria and realistic cohort estimations are crucial for successful study execution. The increasing availability of routine clinical data in electronic health records (EHRs) provides the opportunity to also support patient recruitment during the prescreening phase. While solutions for electronic recruitment support have been published, to our knowledge, no method for the prioritization of eligibility criteria in this context has been explored.

Methods   In the context of the Electronic Health Records for Clinical Research (EHR4CR) project, we examined the eligibility criteria of the KATHERINE trial. Criteria were extracted from the study protocol, deduplicated, and decomposed. A paper chart review and data warehouse query were executed to retrieve clinical data for the resulting set of simplified criteria separately from both sources. Criteria were scored according to disease specificity, data availability, and discriminatory power based on their content and the clinical dataset.

Results   The study protocol contained 35 eligibility criteria, which after simplification yielded 70 atomic criteria. For a cohort of 106 patients with breast cancer and neoadjuvant treatment, 47.9% of data elements were captured through paper chart review, with the data warehouse query yielding 26.9% of data elements. Score application resulted in a prioritized subset of 17 criteria, which yielded a sensitivity of 1.00 and specificity 0.57 on EHR data (paper charts, 1.00 and 0.80) compared with actual recruitment in the trial.

Conclusion   It is possible to prioritize clinical trial eligibility criteria based on real-world data to optimize prescreening of patients on a selected subset of relevant and available criteria and reduce implementation efforts for recruitment support. The performance could be further improved by increasing EHR data coverage.

Keywords
► electronic health records and systems
► data warehousing and data marts
► secondary use
► clinical trial
► recruitment

Background and Significance

Randomized clinical trials are a key component of medical research. Over 1,000,000 trials were performed since 1948.[1] They are regarded as gold standard to test new therapies and diagnostic techniques.[2,3(pp14–15)] Many clinical trials cannot be conducted as planned though.[4,5] Slow recruitment and/or missed target cohort sizes often result in delays, cost overruns, and even cancellation of clinical trials. With rising average costs of €0.8 billion in 2010[6] and €1.9 billion in 2016[7] for research, development and regulatory approval of a new active substance, each failure can be a huge burden for the executing company or academic institution. Even small amendments of the protocol can lead to costs of thousands of euros and delay the trial by requiring ethics approvals for the amendment and implementing the changes in the participating trial centers, as well as delayed time-to-market.[4,8] A good study design with a realistically estimated cohort size is essential to prevent these issues. The increasing availability of structured patient data from electronic health records (EHRs) provides new opportunities toward achieving these goals.

Benefits and Limitations of Real-World Data Use

The reuse of data acquired during routine care in EHRs has been shown to improve both the correct estimation of cohort sizes, as well as the recruitment of study subjects.[9–11] The reported benefits include a simplified and better-targeted identification of recruitment candidates,[10,11] higher rates of accrual,[12–14] as well as time savings, for the process of patient recruitment.[9,11,14] It has further been shown that secondary use of EHR data can prevent repeated data reentry and improve data quality and cost-effectiveness of research.[13,15]

Several limitations regarding secondary use of routine data have been reported: documentation for routine care and billing purposes may introduce selection biases, and their quality and comprehensiveness may not be sufficient for research purposes.[16,17] Patients often visit several health providers, leading to fragmented EHRs.[17] The use of different terminologies (or the lack thereof) complicates and, in some cases, precludes the merging of data from different sources both within an organization or across institutional borders.[18–21]

Availability of Real-World Data

Several authors have examined how eligibility criteria from clinical trials overlap with data items available in EHR systems. Ateya et al[22] decomposed eligibility criteria from 228 studies taken from a U.K. trial repository and used expert classification to determine whether related EHR data elements could likely be used; actual EHR data availability was not assessed. While they found that 74% of the criteria could likely be determined from EHR data, they also noted that EHR queries on their own would be insufficient to determine recruitment and should be seen as a tool to preselect patient cohorts for further manual screening. Köpcke et al additionally assessed the actual availability of data items related to 15-investigator-initiated trials in the EHRs of participating German hospitals and determined that on average, only 35% of criteria were available, as well as documented.[23] Doods et al extracted inventories of commonly used eligibility criteria for feasibility and recruitment from pharmaceutical trials in the Electronic Health Records for Clinical Research (EHR4CR) project and examined the availability of corresponding data elements at participating university hospitals.[24,25] While demographics, diagnosis, and procedure codes and a majority of laboratory findings were highly available, most items from medical history, as well as scores and classifications were rarely present. Löbe et al discussed the consequences of this limited coverage:[26] patient cohorts based on a limited set of electronically available eligibility criteria may overestimate the recruitable population by including false positives that need to be eliminated by manual examination of (paper) patient charts.

Comprehensibility of Eligibility Criteria

Successful implementation of electronic recruitment support also depends on the quality and computability of the eligibility criteria as they are defined in the study protocols. Several publications have examined deficiencies of the current process of defining eligibility criteria: clinical researchers often are not involved in clinical care and documentation and may not have experience on whether certain items (e.g., "able to swallow tablets" or "good health") are routinely captured.[27] Also, a lack of precise understanding regarding etiologies and comorbidities and their relevance to patient eligibility has been observed.[28] On the one hand, a focus on the principal (study) diagnosis may lead to overestimating the size of the target cohort as possible exclusion criteria may not be taken into account.[26] Researchers need to define cohorts of similar probands to ensure that study results depend on the items of interest and no other confounding factors (confusion bias).[2] On the other hand, they need to exclude probands facing disproportionate risks by participating in the study. Researchers often include prior experience in the definition of eligibility criteria,[28] which can be subjective and unsystematic.[29] Post hoc, the detailed reasoning applied during selection of eligibility criteria can often not be reconstructed which may complicate the detection and correction of problems regarding patient enrollment.[28]

Ross et al examined a set of 1,000 trials randomly extracted from ClinicalTrials.gov and assessed the comprehensibility (containing an interpretable criterion), selectiveness (actually affecting candidate selection) and complexity (atomic versus combined criteria) of their eligibility criteria.[30] They found 7% of the criteria to be incomprehensible or nonselective. Of the remaining 932 criteria, only 15% were found to be simple criteria containing discrete clinical concepts in single phrases or quantitative comparisons. The other 85% were complex criteria that contained multiple concepts, temporal constraints, or complex comparisons that would require decomposition into distinct statements, as well as criteria that would require clinical judgement, or information beyond the eligibility criteria (e.g., from the study protocol). Girardeau et al applied this classification to three studies within the EHR4CR project and also assessed EHR availability and computability of criteria.[31] They noted the influence of missing data on reducing the sensitivity (when regarding inclusion criteria) and specificity (regarding exclusion criteria) of the queries.

Categorization of Eligibility Criteria

Wang et al implemented a similar approach toward the categorization of eligibility criteria[27] by assessing the effort of electronic implementation:

- "Easy" (supporting fully automated queries).
- "Mixed" (supporting automated queries with subsequent manual checks).
- "Hard" (requiring fully manual retrieval).
- "Impossible" (not routinely documented in the EHR).

Using ClinicalTrials.gov, they found 292 individual criteria in a convenience sample of 20 studies, removed duplicate and redundant criteria, and categorized them by six independent observers from two separate research institutions, leading to the following groups:

- "Easy": laboratory findings, diagnoses or procedures.
- "Mixed": diagnoses with modifiers (e.g., "severe cardiovascular disease [defined by NYHA $\geq$ 3]," "active or untreated latent tuberculosis [TB]").
- "Hard": criteria usually found in narrative clinical notes (e.g., "females who are breastfeeding," "eastern cooperative oncology group [ECOG] performance status of 0 or 1").
- "Impossible": temporally related or generally undocumented criteria (e.g., "presenting within timeframe for intravenous tPA treatment approved by local regulatory authorities but no more than 4.5 hours from onset of symptoms," "facial hair" or "good health").

Role in Patient Recruitment

Trinczek et al proposed a generic software architecture for patient recruitment systems (PRS)[32] consisting of five modules: trial administration module, notification module, patient data module, query module, and screening list module.
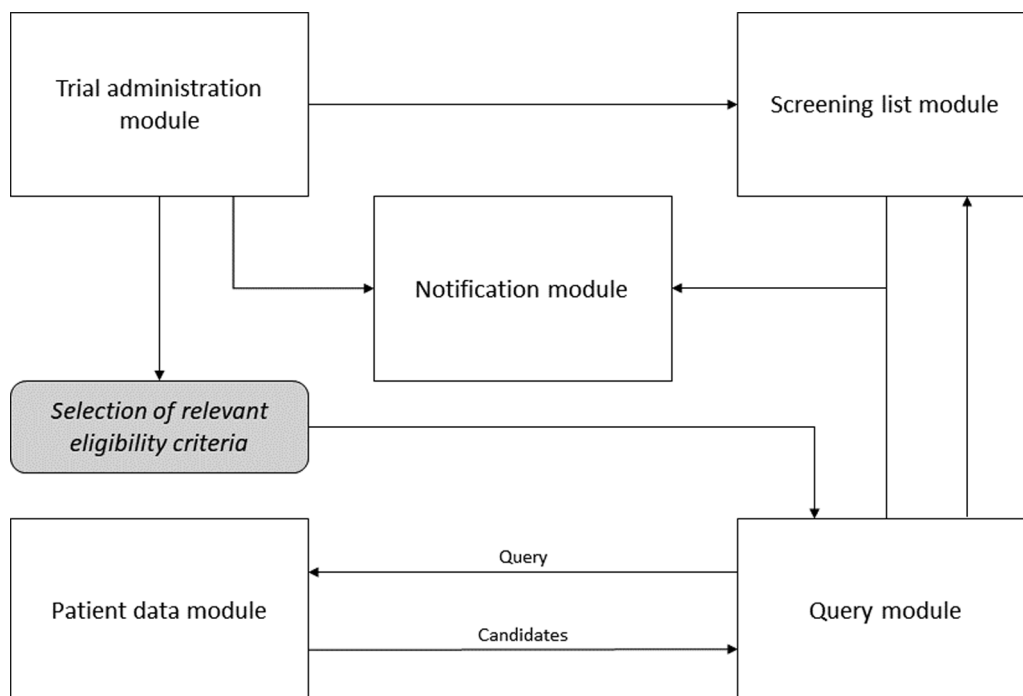
In the query module, eligibility criteria transferred from the trial administration module are converted to executable queries, which are then applied to the patient data module. The authors also posited that the selection of the eligibility criteria to implement electronically is crucial. We have added this activity as a separate step in the recruitment architecture diagram proposed by Trinczek et al between the trial administration and query modules (►Fig. 1).

Cuggia et al have also emphasized the critical nature of clearly formulated and interpretable eligibility criteria.[2] They have posited that prescreening is a crucial step of the recruitment process. Prescreening is described as an initial selection of potential candidates from a proband population based on a subset of prioritized eligibility criteria. Only subsequently, candidates are screened against the full set of criteria. Even though Cuggia et al reviewed 28 recruitment support publications, none of the eight, which were presented in detail, described a process of how the prioritized subset was selected. In this paper, we propose a stepwise approach toward the selection of such a prioritized subset.

Objectives

In this project, we analyzed the eligibility criteria of a clinical trial with the goal of developing a systematic approach toward identifying a relevant subset of criteria best suited to implement recruitment support, based on availability in the EHR and their discriminatory power. To our knowledge, no systematic approach toward selection of relevant eligibility criteria for recruitment support has been published so far.

The project was performed within the EHR4CR project, a European Union Innovative Medicines Initiative (EU-IMI) funded public–private partnership, which focused on the

Fig. 1 Generic patient recruitment support software architecture (modified after Trinczek et al[32]), supplemented with a criteria selection step.

optimization of clinical trials throughout the feasibility, recruitment, execution, and pharmacovigilance phases.[12]

Methods

Based on trials by the participating pharma companies that were actively recruiting during the project phase, the KATHERINE study was selected for the project (NCT01772472). The KATHERINE study compares the efficacy and safety of trastuzumab–emtansine versus trastuzumab in breast cancer for patients with HER2-positive residual tumors after tumor resection and neoadjuvant therapy.

Permission to carry out the study was granted by the ethics board of the Medical Faculty of the Friedrich-Alexander University Erlangen-Nürnberg (247_14Bc). All eligibility criteria were extracted from the study protocol and not from the simplified version available at ClinicalTrials.gov. They were classified according to the comprehensibility, selectiveness, and complexity aspects by Ross et al[30] and compared with the previously published EHR4CR data inventories.[15,24,25] The eligibility criteria were then simplified to allow algorithmic implementation. To prepare the eligibility criteria for electronic execution, we classified and refined them according to Ross et al[30] in the following stepwise approach:

- Identification of incomprehensible or nonselective criteria which were eliminated from further processing.
- Identification of duplicate criteria (also covering inclusion criteria which were duplicated as inversely formulated exclusion criteria) which were reduced to a single instance.
- Identification of noncomputable criteria (e.g., requiring physician interpretation) which were eliminated from further processing.
- Identification of complex criteria (i.e., containing several attributes in a single clause) which were decomposed into simple criteria.

Based on the respective sections of the study protocol, criteria were also tagged as disease specific versus non-disease-specific.

Eligibility criteria were then matched with data elements from the local clinical data warehouse at Erlangen University Hospital, a tertiary-care academic site with 1,394 beds. While the hospital offers specialized outpatient clinics, ambulatory care in Germany is covered primarily through general practitioners not affiliated with the hospitals. Data warehouse content thus relates mostly to inpatient care. A preselection of relevant patient identifiers was extracted from the local tumor documentation system (GTDS, Gießen University) based on documented breast cancer (ICD code "C50") and neoadjuvant therapy (T-stage "y" in the TNM classification of malignant tumors) during the time period from March 25, 2013–October 27, 2014. Available data elements for the selected cohort were exported from an i2b2 platform.[33] At the same time, a paper chart review was performed for the same cohort to manually extract all documented eligibility data elements.

The availability of all eligibility data elements was calculated both for the data warehouse extract, as well as the data manually extracted by chart review. Data element availability was compared with the EHR4CR feasibility data inventory.[25]

The refined set of eligibility criteria for electronic execution consisted only of comprehensible, selective and simple criteria. Data availability was computed for all criteria in the set based both on the dataset generated from chart review, as well as from the data warehouse. To quantify discriminatory power of criteria, an isolated inclusion or exclusion result was determined for each available value in both datasets. Missing values were considered to be "neutral" in the sense of resulting neither in an inclusion or exclusion. A comparison was performed between the eligibility results of both data sources. In case of discrepancies, the reasons were determined and documented based on reviewing the patient chart and raw data in the clinical data warehouse. Also, the specificity of the combined disease-specific and nondisease-specific criteria were calculated, respectively. In a further step, we applied a score to select criteria most suitable for electronic execution, based on the following components:
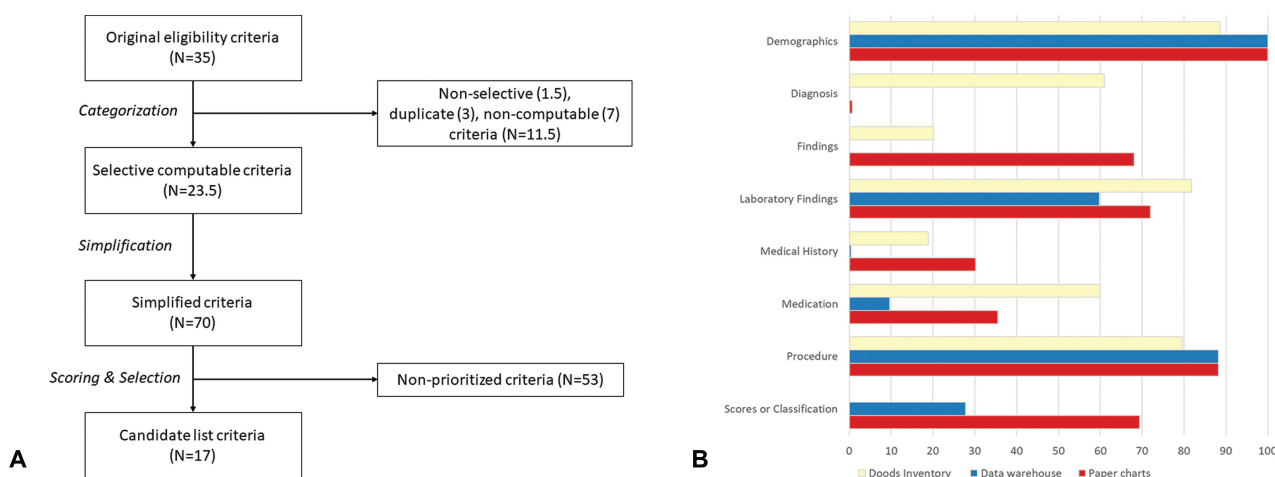
- Disease specificity: criteria listed in the "disease-specific eligibility" sections of the protocol received a point.
- Data availability: criteria which had data available from the paper chart or data warehouse received a point.
- Discriminatory power: criteria which were discriminatory (i.e., with available data leading to patient exclusion) received a point.

The score was added for each criterion, and a threshold of 2 was defined for inclusion into the final set of eligibility criteria for electronic execution.

The screening list was obtained from the principal investigator to provide the patients actually included in the study. The inclusion list was used to calculate sensitivity and specificity for the selected eligibility criteria. Classification and scoring of eligibility criteria, as well as the paper chart reviews, were performed by a fifth-year medical student (G. M.) and vetted by a medical doctor (T.G.).

Results

The selection and refinement of eligibility criteria is shown in ►Fig. 2A, B. All 35 original eligibility criteria were determined to be comprehensible, with 1.5 criteria being nonselective and 33.5 selective according to the taxonomy described by Ross et al[30] ("fractional" criteria are given when a complex criterion contains multiple simple criteria with different classifications). Of these, 10 criteria were classified as "simple" and 23.5 as "complex." After the removal of duplicates (3) and noncomputables (7) and the decomposition of the complex criteria into simple components, a total of 70 individual criteria resulted (►Table 1; ►Supplementary Table S1 [available in the online version]; for the detailed list). Data elements for 53 criteria (75.7%) were available in the local clinical data warehouse. 47 (67.1%) items were included in the EHR4CR trial feasibility inventory,[25] 48 items (68.6%) were part of the EHR4CR recruitment inventory,[24] and 47 items (67,1%) were part of the EHR4CR trial execution inventory.[15]

Fig. 2 (A) Flowsheet describing the selection and refinement of eligibility criteria, starting from the original criteria as given in the study protocol. After a categorization step, non-selective, duplicate and non-computable criteria are removed. Simplification of the remaining criteria leads to a set of 70 criteria, of which 17 are selected after scoring (see ►Fig. 2B for details). (B) Flowsheet describing the scoring and selecting criteria to achieve a prioritized subset. Starting from a cohort of breast cancer patients with neoadjuvant first-line therapy, data for the full set of simplified eligibility criteria is extracted in parallel from a data warehouse query and a paper chart review. The eligibility impact of each available data element is annotated to derive a master data table giving the availability and eligibility impact of each data element for all patients from both sources. A score is calculated based on the disease-specificity, data availability and discriminatory power of each criterion.

Table 1 Composition of eligibility criteria; "fractional" criteria are given when a complex criterion contains several simple criteria with different classifications

Original set	35 criteria n (%)
• Nonselective criteria	1.5 (4.3)
• Selective criteria	33.5 (95.7)
■ Simple	10 (28.6)
■ Complex	23.5 (67.1)
Categorization of selective criteria	33.5
• Duplicate criteria	3 (9.0)
• Noncomputable criteria	7 (20.9)
• Nonduplicate, computable criteria	23.5 (70.1)
Decomposed (simplified) set	70 criteria
• Available in local data warehouse	53 (75.7)
• Present in EHR4CR feasibility criteria inventory[25]	47 (67.1)
• Present in EHR4CR recruitment criteria inventory[24]	48 (68.6)
• Present in EHR4CR trial execution inventory[15]	47 (67.1)

Abbreviation: EHR4CR, Electronic Health Records for Clinical Research.

The preselection of relevant patients from the GTDS tumor documentation system yielded 115 patient identifiers. Of these, 106 (92%) paper charts could be retrieved during the study period, whereas 9 (8%) charts were unavailable due to clinical use. These patients were excluded from the project. Manual chart review to extract the computable data items took 32.5 hours (18.8 minutes on average). It took 15.5 hours to determine availability of data items and extract them from the clinical data warehouse. Out of a total of 7,420 possible data elements (70 items for 106 patients), 3,551 (47.9%) were available from the paper charts and 1,995 (26.9%) from the data warehouse.
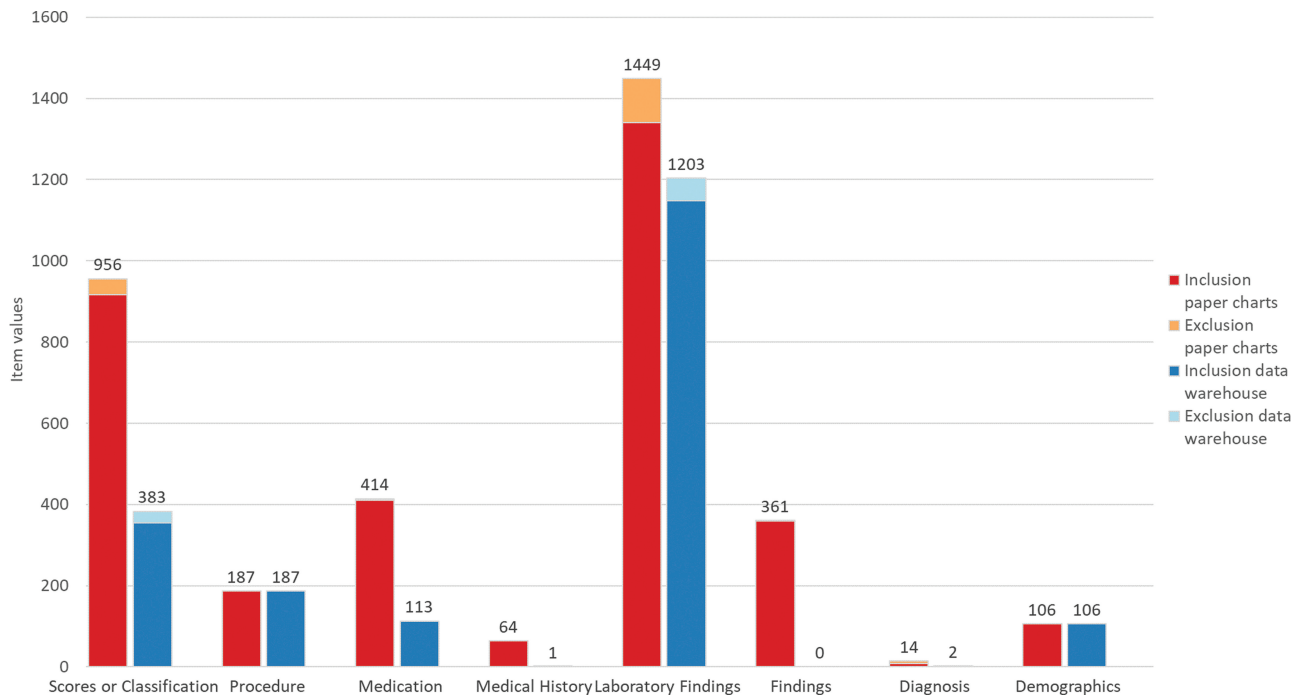
►Fig. 3 shows the data availability for the paper chart review and data warehouse, aggregated by the groups from the EHR4CR recruitment inventory and compared with the availability listed there[24] (►Supplementary Figs. S1 and S2 [available in the online version] for a detailed breakdown). ►Table 2 shows data availability for these groups in the local data warehouse compared with the EHR4CR feasibility and trial execution inventories. Eligibility results were determined for each individual data item for both sources (►Fig. 4). Results were concordant for 1,930 data elements (99.7%) and differed in five cases (0.3%). The reasons for these differences were determined and are given in ►Table 3.

Specificity of the combined disease-specific criteria was 0.55 and 0.13 for the combined nondisease-specific criteria. Application of the scoring system led to 2 criteria receiving the maximum score of 3 points, 15 criteria with 2 points, 37 criteria with 1, and 16 criteria with 0 points. Based on the cut-off at 2 points, a set of 17 criteria were selected for electronic execution (►Table 4).

Application of the criteria against the data warehouse dataset and screening list yielded a sensitivity of 1.00, a specificity of 0.57, a positive predictive value of 0.10, and a negative predictive value of 1.00 (►Table 5).
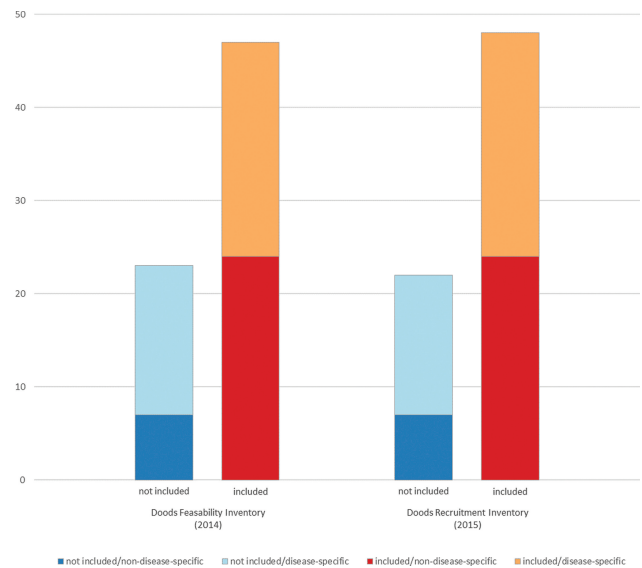
Discussion

Prescreening has been described as an essential but challenging step within the recruitment process, facilitating an initial selection of potentially recruitable patients from a base population,[2,32,34] based on a limited set of criteria available from electronic sources and followed up by in-depth manual review of the candidates against the full set of

Fig. 3 Data availability for the paper chart review and data warehouse, aggregated by the groups from the EHR4CR recruitment inventory and compared with the availability listed there[24] (►Table 2 for accompanying data). EHR4CR, Electronic Health Records for Clinical Research.

Table 2 Accompanying data table for ►Fig. 3 showing data availability from paper chart and data warehouse and the Inventory by Doods et al[25]; availability from the inventory was averaged across the nine participating hospitals

Criteria group	Paper chart review (%)	Data warehouse query (%)	Inventory by Doods et al (%)
Demographics	100.0	100.0	88.6
Medical history	30.2	0.5	18.9
Diagnosis	0.8	0.1	61.0
Procedure	88.2	88.2	79.6
Findings	68.1	0.0	20.2
Laboratory findings	71.9	59.7	81.8
Medication	35.5	9.7	60.0
Scores or classification	69.4	27.8	0.0

Fig. 4 Isolated eligibility results of individual eligibility criteria grouped by inventory by Doods et al,[24] paper chart review and the data warehouse query.

eligibility criteria. This aligns with the expectations of potential users of electronic recruitment support in the sense that a PRS would not be expected to provide a definite list of patients to be recruited, but rather a relevant preselection for further manual inspection.[32]

Applying the categorization proposed by Ross et al[30] to KATHERINE study, the composition of eligibility criteria was similar to that reported by Ross et al: 95.7% of criteria were comprehensible (Ross et al: 93.2%), and among those 72% were complex and 28% were simple criteria (Ross et al: 85/15%). We

applied a stepwise process of categorizing, pruning and electronic implementation of criteria that allowed us to reduce the effort required for setting up electronic recruitment support. We support the recommendation by Ross et al and van Spall et al[30,35] that eligibility criteria should be formulated in a comprehensible, selective, and simple manner to provide a concise set of consistent criteria suitable for electronic implementation. In a more recent project, Zhang et al analyzed eligibility criteria from 77 Hepatitis C-Virus (HCV) trials in 2018,[36] finding 85% of criteria to be computable and proposed

Table 3 Mismatching criteria between paper charts and data warehouse

Criterion	Affected patients	Comment
TNM (post-op)—T/TNM (post-op)—R	3	The data warehouse query returned data for earlier pathology findings which fell into the relevant time window
laboratory value—hemoglobin	1	The data warehouse query returned an earlier finding
anticancer drug—epirubicin, cycles	1	The paper charts contained more detailed information including the epirubicin cycles (leading to exclusion), whereas the data warehouse contained only the general information, that epirubicin was given

Table 4 Number of criteria in relation to the score as well as sensitivity and specificity calculated for each combination both from the data warehouse query and the paper chart review

	Number of criteria	Data warehouse query		Paper chart review	
		Sens.	Spec.	Sens.	Spec.
Score 3	2	1.000	0.554	1.000	0.772
Score 2 or above	17	1.000	0.574	1.000	0.802
Score 1 or above	54	1.000	0.574	1.000	0.811

Table 5 Contingency table of inclusion and exclusion of patients comparing electronic execution (PRS recommendation) of eligibility criteria from the data warehouse versus the trial screening list

		Trial screening list	
		Inclusion	Exclusion
PRS recommendation	Inclusion	5	43
	Exclusion	0	58

Note: Sensitivity: 1.00, specificity: 0.57, positive predictive value: 0.10, negative predictive value: 1.00.

a classification of eligibility criteria related to their ontology-based operationalization. Combining simplification based on Ross et al and prioritization as described in this paper with an ontology-based implementation could improve generalizability, for example, regarding application across different terminologies and granularities of clinical data.

Even though previous publications have stated the importance of prioritizing relevant eligibility criteria for the pre-screening step,[2] according to our knowledge, no concrete process has been published regarding how to select this relevant subset. We additionally applied a scoring system to select a subset of criteria most relevant for building candidate lists for electronic recruitment support and evaluated it based on a comparison with the patients actually recruited into the trial. We performed a preselection of relevant patients based on core eligibility criteria (breast cancer and neoadjuvant therapy). This step provided us with a dataset that we could then analyze in relation to the availability and value distribution of the remaining eligibility criteria specific to that selected cohort. We chose to prioritize disease-specific criteria as their combined specificity was higher (0.55) than the combined nondisease-specific criteria (0.13). Additionally, we prioritized criteria for which the available data in the preselected cohort was nonuniform regarding inclusion or exclusion (i.e., data elements which did not either include or exclude all patients homogeneously) to ensure that only criteria leading to a discrimination within the preselected cohort were used. Finally, data elements with no available data were downranked, as an implementation of the data availability categorization proposed by Wang et al.[27] The cut-off was set at a score value of 2, as this set of criteria achieved a higher specificity than a score of 3. Based on the paper chart review, the effect was even higher. The cut-off was not set at score value 1, as this would have been almost identical to using the full set of criteria (54 out of 70).

The resulting sensitivity and specificity show that the cohort derived from our prioritized subset of eligibility criteria is larger than the set of actually recruited patients. The set includes false-positive patients but no false negatives. While this ensures that no potential candidate was excluded during the prescreening step, the false-positive candidates require additional manual inspection. This matches the observation regarding false positives as a result of limited data availability from Löbe et al.[26]

Comparison of the eligibility criteria of the KATHERINE study with the inventories published by Doods et al[24,25] (►Fig. 3) showed only partial coverage, due to the structure of the inventories. While some attributes are listed with generic labels (e.g., "verbatim drug name" for medications), the actual trial eligibility criteria referred to specific substances (e.g. "Doxorubicin"). The inventory structure could be considered inconsistent with regard to the fact that laboratory findings are not grouped but given individually (e.g., "total cholesterol in serum"). Given diagnoses yielded a very low availability both in the data warehouse, as well as chart review, in comparison to the inventory. The study protocol referred to a set of specific diseases as exclusion criteria (which were rare in the cohort), whereas in the inventory covered the presence of any diagnosis in the dataset. We also noted that disease-specific criteria (e.g., number of chemotherapy cycles) are underrepresented in the inventories (►**Supplementary Figs. S1** and **S2**; available in the online version). As the inventories were generated by analyzing the frequency of criteria across a large set of

studies, it follows that criteria relevant across several diseases shared higher frequencies whereas disease-specific criteria would not reach the required threshold for inclusion into the inventories. In our dataset, disease-specific criteria had a higher relevance toward the selection of a prioritized subset of criteria. It should be considered to extend the inventories with disease group–specific modules.

Löbe et al, Trinczek et al, and Zhang et al[26,32,36] noted that clinical trial candidate identification and screening for recruitment currently are very time-consuming manual tasks. In our project, manual chart review of the full set of eligibility criteria in a base population of 106 patients took 32.5 hours, whereas the time spent for constructing and executing the data warehouse query was 15.5 hours. Since query implementation efforts relate only to the number of criteria implemented, whereas manual chart review relates to the size of the cohort, the potential gains of electronic execution should increase with the size of the base population.

Averitt et al compared cohort compositions of four landmark randomized controlled trials (RCTs) with cohorts derived from routine clinical datasets[37] and found that even though identical eligibility criteria were rigorously applied, baseline summary statistics varied between published results and EHR-derived datasets, suggesting heterogeneity of treatment effects (HTE) and putting replicability, as well as medical applicability, of RCT results on real-world cohorts into doubt. Among other measures, Averitt et al propose to implement a more structured, codified documentation of eligibility criteria to enhance replicability. While electronic recruitment support could help to standardize application of eligibility criteria, data availability, and selection of implementable/prioritized criteria could introduce biases of their own.

Limitations

The KATHERINE trial chosen for this project has very narrow eligibility criteria, resulting in a very small percentage of actually recruitable patients within the available base population, also contributing to the low–positive predictive value of 0.1%. Potential patients which matched the eligibility criteria but may have declined to participate in the study were not taken into account. Also, the project was performed only at a single academic hospital with data availability from the routine care process limited mostly to inpatient care. This could negatively impact the applicability of the results on trials with broader criteria and/or other types of hospitals. The selection of the trial to be used in the project was constrained by the scope and collaborating partners in the relevant EHR4CR work package. The analysis of data availability was performed not against the full patient population of the hospital, but against a preselected cohort matching basic criteria (breast cancer and neoadjuvant treatment) determined from a separate documentation platform not included in the data warehouse. A full paper chart review would not have been feasible on the full patient population. This mandatory preselection step is in fact an integral part of the proposed approach for determining the prioritized subset of eligibility criteria to implement for patient recruitment support. Beyond review of the cited literature, no specific training was applied for the staff carrying out the classification and scoring of the eligibility criteria. Subjectivity cannot be ruled out for some of the classification decisions (e.g., comprehensibility). Whether the selected cut-off of score value 2 for inclusion of criteria can be generalized to other trials needs to be confirmed. In this project, only structured data elements in the clinical data warehouse were examined, and specifically no natural language processing (NLP) approaches were performed to extract additional data from narrative text (e.g., discharge letters). While NLP is increasingly being applied on English-language datasets, it is not yet broadly implemented for German-language EHRs.

Conclusion

Patient recruitment support for clinical trials based on electronic health records is a topic of continuing interest. The prescreening step has been identified as the focal point of establishing efficient recruitment support, yet no systematic process for identifying a prioritized subset of eligibility criteria has yet been published. Our proposed approach facilitates a data-driven selection of items based on their relevance to the trial, the actual availability of data in the EHR and the resulting discriminatory power of the chosen criteria. Apart from streamlining implementation of electronic recruitment support, the approach could also be leveraged during the protocol design, as well as site selection/feasibility phase. While the increasing availability of structured EHR data provides an opportunity for secondary use in the context of clinical trials, the quality of eligibility criteria in study protocols with regard to their consistency and interpretability remains an important issue that needs to be addressed. Annotation with standardized terminologies, inventories of commonly used criteria (including disease-specific aspects) and possibly even reusable databases of criteria[38] could be leveraged to simplify the implementation of electronic recruitment support.

With the current implementation of large-scale secondary use infrastructures, like the German Medical Informatics Initiative (MII)[39] or the Swiss Personalized Health Network (SPHN),[40] harmonized platforms are currently becoming available that will also facilitate patient recruitment support. Within the MII, the Medical Informatics in Research and Care in University Medicine (MIRACUM) consortium pursues patient recruitment support as a primary use case,[41] providing an infrastructure for a multicentric implementation and evaluation of the approach presented in this paper.

Clinical Relevance Statement

Inclusion of patients in clinical trials is an integral part of, but not limited to, academic medicine and can contribute to certification criteria (e.g., in comprehensive cancer centers). Leveraging real-world data to support the recruitment process addresses the need of hospitals to optimize the execution of clinical trials.

Multiple Choice Questions

1. How should eligibility criteria for clinical trials be formulated?
   a. Complex and nonselective
   b. Selective and atomic
   c. Redundant and computable
   d. Verbose and machine readable

   Correct Answer: The correct answer is option b. Eligibility criteria for clinical trials should be formulated to be selective (i.e., contain criteria that can be applied to derive an eligible subset of probands) and atomic in the sense of limiting each criterion to a single attribute. Complex criteria (containing several attributes) should be decomposed into sets of atomic criteria. Redundancy should be avoided, including cases in which a criterion appears both in the inclusion section (e.g., patients with M0 status), as well as in a negated form, in the exclusion section (e.g., patients with M1 status). While it is desirable to have computable eligibility criteria (i.e., being able to electronically derive from EHR data), in many cases criteria need physician interpretation (e.g., whether a patient can be expected to comply to the study protocol). Eligibility criteria should be formulated concisely, avoiding verbosity.

2. How can real-world data (RWD) support the execution of clinical trials?
   a. RWD fully automates recruitment and execution of clinical trials
   b. RWD obviates physician interpretation of eligibility criteria
   c. RWD fully covers all attributes used for determining clinical trial eligibility
   d. RWD can support adequate cohort size estimation and be used to select recruitment candidates

   Correct Answer: The correct answer is option d. RWD data typically covers only a subset of attributes required for determining clinical trial eligibility, excluding, for example, data not documented electronically, requiring physician interpretation or not documented within a relevant timeframe. Applying a relevant subset of available, selective, and prioritized data elements can be leveraged to achieve an adequate estimation of cohort size and select candidates for prescreening. Final decisions toward inclusion or exclusion into a trial need to be made by qualified personnel based not only on electronically available data but also on data available from the paper chart or attributes acquired during the screening process.

Note

The present work was performed in fulfillment of the requirements for obtaining the degree "Dr. med." from the Friedrich-Alexander University Erlangen-Nürnberg (FAU).

Protection of Human and Animal Subjects

The project was performed in compliance with the World Medical Association Declaration of Helsinki on Ethical Principles for Medical Research Involving Human Subjects, and was reviewed by the ethics board of the Medical Faculty of the University of Erlangen Nuremberg (247_14Bc).

Conflict of Interest

T.G. reports grants from European Commission during the conduct of the study. G.M. reports grants from European Commission, during the conduct of the study. H.U.P. reports grants from European Commission, during the conduct of the study.

References

1 Victor N. Registration of Clinical Studies from View of Ethics Committees (german language). Deutsches Ärzteblatt 2004;101 (30):A-2111/B-1763/C-1695
2 Cuggia M, Besana P, Glasspool D. Comparing semi-automatic systems for recruitment of patients to clinical trials. Int J Med Inform 2011;80(06):371–388
3 Schumacher M, Schulgen G. Controlled clinical trials - an introduction (german language). In: Methodology of clinical studies. 2008:1–19
4 Kalra D, Schmidt A, Potts H, Dupont D, Sundgren M, de Moor G. Case report from the EHR4CR project—A European Survey on Electronic Health Records Systems for Clinical Research. iHealth Connections 2011:108–113
5 Prescott RJ, Counsell CE, Gillespie WJ, et al. Factors that limit the quality, number and progress of randomised controlled trials. 1999;3(20):1–143
6 Fink T, Wicke D. Clinical trial - challenge with significant impact: (german language). Biotechnologie '10, '11 - Kapital, Markt, Wirtschaft. 2010. Accessed November 22, 2014
7 EFPIA. The Pharmaceutical Industry in Figures: Key Data 2018. Available at: https://www.efpia.eu/media/361960/efpia-pharmafigures2018_v07-hq.pdf. Accessed October 21, 2020
8 Getz KA, Stergiopoulos S, Short M, et al. The impact of protocol amendments on clinical trial performance and cost. Ther Innov Regul Sci 2016;50(04):436–441
9 Liu K, Acharya A, Alai S, Schleyer TK. Using electronic dental record data for research: a data-mapping study. J Dent Res 2013; 92(7, suppl)90S–96S
10 Embi PJ, Jain A, Clark J, Harris CM. Development of an electronic health record-based clinical trial alert system to enhance recruitment at the point of care. AMIA Annu Symp Proc 2005; 2005:231–235
11 Köpcke F, Kraus S, Scholler A, et al. Secondary use of routinely collected patient data in a clinical trial: an evaluation of the effects on patient recruitment and data acquisition. Int J Med Inform 2013;82(03):185–192
12 De Moor G, Sundgren M, Kalra D, et al. Using electronic health records for clinical research: the case of the EHR4CR project. J Biomed Inform 2015;53:162–173
13 Bruland P, Forster C, Breil B, Ständer S, Dugas M, Fritz F. Does single-source create an added value? Evaluating the impact of introducing x4T into the clinical routine on workflow modifications, data quality and cost-benefit. Int J Med Inform 2014;83(12): 915–928
14 Dugas M, Lange M, Müller-Tidow C, Kirchhof P, Prokosch H-U. Routine data from hospital information systems can support

patient recruitment for clinical studies. Clin Trials 2010;7(02): 183–189

15 Bruland P, McGilchrist M, Zapletal E, et al. Common data elements for secondary use of electronic health record data for clinical trial execution and serious adverse event reporting. BMC Med Res Methodol 2016;16(01):159

16 Hersh WR, Weiner MG, Embi PJ, et al. Caveats for the use of operational electronic health record data in comparative effectiveness research. Med Care 2013;51(08, Suppl 3):S30–S37

17 Weiner MG, Embi PJ. Toward reuse of clinical data for research and quality improvement: the end of the beginning? Ann Intern Med 2009;151(05):359–360

18 Weng C, Tu SW, Sim I, Richesson R. Formal representation of eligibility criteria: a literature review. J Biomed Inform 2010;43 (03):451–467

19 Blaisure J, Ceusters W. Business rules to improve secondary data use of electronic healthcare systems. Stud Health Technol Inform 2017;235:303–307

20 Bache R, Taweel A, Miles S, Delaney BC. An eligibility criteria query language for heterogeneous data warehouses. Methods Inf Med 2015;54(01):41–44

21 Ash JS, Anderson NR, Tarczy-Hornoch P. People and organizational issues in research systems implementation. J Am Med Inform Assoc 2008;15(03):283–289

22 Ateya MB, Delaney BC, Speedie SM. The value of structured data elements from electronic health records for identifying subjects for primary care clinical trials. BMC Med Inform Decis Mak 2016;16:1

23 Köpcke F, Trinczek B, Majeed RW, et al. Evaluation of data completeness in the electronic health record for the purpose of patient recruitment into clinical trials: a retrospective analysis of element presence. BMC Med Inform Decis Mak 2013;13:37

24 Doods J, Lafitte C, Ulliac-Sagnes N, et al. A European inventory of data elements for patient recruitment. Stud Health Technol Inform 2015;210:506–510

25 Doods J, Botteri F, Dugas M, Fritz FEHR4CR WP7. A European inventory of common electronic health record data elements for clinical trial feasibility. Trials 2014;15:18

26 Löbe M, Stäubert S, Goldberg C, Haffner I, Winter A. Towards phenotyping of clinical trial eligibility criteria. Stud Health Technol Inform 2018;248:293–299

27 Wang AY, Lancaster WJ, Wyatt MC, Rasmussen LV, Fort DG, Cimino JJ. Classifying clinical trial eligibility criteria to facilitate phased cohort identification using clinical data repositories. AMIA Annu Symp Proc 2018;2017:1754–1763

28 Weng C. Optimizing clinical research participant selection with informatics. Trends Pharmacol Sci 2015;36(11):706–709

29 Rubin DL, Gennari J, Musen MA. Knowledge representation and tool support for critiquing clinical trial protocols. Proc AMIA Symp 2000:724–728

30 Ross J, Tu S, Carini S, Sim I. Analysis of eligibility criteria complexity in clinical trials. Summit On Translat Bioinforma 2010; 2010:46–50

31 Girardeau Y, Doods J, Zapletal E, et al. Leveraging the EHR4CR platform to support patient inclusion in academic studies: challenges and lessons learned. BMC Med Res Methodol 2017;17(01): 36

32 Trinczek B, Köpcke F, Leusch T, et al. Design and multicentric implementation of a generic software architecture for patient recruitment systems re-using existing HIS tools and routine patient data. Appl Clin Inform 2014;5(01):264–283

33 Murphy SN, Weber G, Mendis M, et al. Serving the enterprise and beyond with informatics for integrating biology and the bedside (i2b2). J Am Med Inform Assoc 2010;17(02):124–130

34 Schreiweis B, Bergh B. Requirements for a patient recruitment system. Stud Health Technol Inform 2015;210:521–525

35 Van Spall HGC, Toren A, Kiss A, Fowler RA. Eligibility criteria of randomized controlled trials published in high-impact general medical journals: a systematic sampling review. JAMA 2007;297 (11):1233–1240

36 Zhang H, He Z, He X, et al. Computable eligibility criteria through ontology-driven data access: a case study of hepatitis C virus trials. AMIA Annu Symp Proc 2018;2018:1601–1610

37 Averitt AJ, Weng C, Ryan P, Perotte A. Translating evidence into practice: eligibility criteria fail to eliminate clinically significant differences between real-world and study populations. NPJ Digit Med 2020;3:67

38 Ash N, Ogunyemi O, Zeng Q, Ohno-Machado L. Finding appropriate clinical trials: evaluating encoded eligibility criteria with incomplete data. Proc AMIA Symp 2001:27–31

39 Gehring S, Eulenfeld R. German medical informatics initiative: unlocking data for research and health care. Methods Inf Med 2018;57(S 01):e46–e49

40 Baillie Gerritsen V, Palagi PM, Durinx C. Bioinformatics on a national scale: an example from Switzerland. Brief Bioinform 2019;20(Suppl 2):361–369

41 Prokosch H-U, Acker T, Bernarding J, et al. MIRACUM: medical informatics in research and care in university medicine. Methods Inf Med 2018;57(S 01):e82–e91