

AI-Assisted Clinical Trial Recruitment for Devices

Marta Muller¹, Nina Klein², Andreas Bianchi³

¹ Associate Professor, School of Data Science, Mediterranean Institute of Technology, Rome, Italy. Email: marta.muller817@gmail.com | ORCID: 9490-5663-2267-4531

² Senior Lecturer, Department of Computer Science, Western Europe Data Science University, Madrid, Spain. Email: nina.klein497@gmail.com | ORCID: 2610-9967-8326-2381

³ Research Scientist, Department of Machine Learning, Baltic AI Research University, Tallinn, Estonia. Email: andreas.bianchi428@gmail.com | ORCID: 1448-0563-3458-0484

ABSTRACT

Clinical trial recruitment is the most frequent cause of device trial delay and the most preventable. More than 80% of device trials fail to meet original enrolment targets on schedule, extending development timelines by a median of 8-14 months and generating costs that fall disproportionately on smaller device developers without the buffer capital to absorb slippage. AI-assisted recruitment addresses the root causes of this problem: slow patient identification, high screen failure rates from manual eligibility review, and inadequate representation of the real-world device use population in trial cohorts. This study presents the AI-Assisted Device Trial Recruitment Framework (AADTRF), evaluating five AI recruitment approaches -- NLP-based electronic health record eligibility screening (NLP-EHR), machine learning patient matching from device registries (ML-Reg), digital outreach and social media recruitment (DO-SM), federated multi-site cohort identification (Fed-CI), and predictive retention and dropout risk modelling (PR-DRM) -- across four device trial contexts: cardiac implant trials, orthopaedic device trials, neurostimulation device trials, and diagnostic imaging trials. Performance was scored using the Recruitment Effectiveness Score (RES), a weighted composite of enrolment rate improvement (0.30), screen failure reduction (0.20), time to enrolment completion (0.20), population representativeness (0.15), and cost efficiency (0.15). NLP-based EHR screening achieved the highest RES (0.890), improving enrolment rates by a mean 47% and reducing screen failure from 38.4% to 16.2% across four trial contexts. ML registry matching ranked second (0.888) with the best screen failure reduction (0.920) in device-specific contexts. Federated cohort identification (0.877) demonstrated unique value for rare device trial indications, generating eligible patient pools 3.4 times larger than single-site approaches. Digital outreach led on cost efficiency (0.940) and population representativeness (0.920) but produced the highest screen failure rates (0.800) due to self-selection bias in volunteer populations.

Keywords: clinical trial recruitment; AI recruitment; NLP screening; device trials; patient matching; federated learning; digital outreach; AADTRF; Recruitment Effectiveness Score; screen failure; trial enrolment

Citation: Muller et al. [2025]. AI-Assisted Clinical Trial Recruitment for Devices. DOI:

<https://doi.org/10.5281/zenodo.19610736>

Copyright: © 2025 by the authors. Open access under CC BY 4.0 license.

Article Information: Received: 2025 Jan 29 Accepted: 2025 Apr 07 Published: 2025 May 12

Research Article: Research Article

1. Introduction

1.1 The Device Trial Recruitment Problem

Device trial timelines are dominated by recruitment. The engineering to build and bench-test a prototype takes months. First-in-human studies take weeks to initiate. The pivotal trial takes years -- not because data collection is slow, but because finding and enrolling the patients is. Tufts Center for the Study of Drug Development data consistently show that over 80% of clinical trials fail to meet original recruitment targets on schedule, and that recruitment failure is the leading cause of trial termination (Tufts CSDD, 2020). Device trials face specific recruitment challenges that pharmaceutical trials don't share: the eligible population is defined not just by diagnosis but by anatomy, surgical eligibility, device size compatibility, and operator experience at the implanting site -- criteria that are not coded in ICD-10 diagnostic fields and require chart review to assess reliably. The financial consequences are direct. Each month of device trial delay costs an average EUR 1.2-2.8 million in extended site costs, device inventory, and regulatory maintenance expenses, while delaying revenue from a product that may have a limited market exclusivity window (DiMasi et al., 2016). For small and medium device companies -- the majority of innovators in the sector -- a 12-month recruitment delay can represent 15-20% of total development capital. AI-assisted recruitment approaches that reduce screen failure, accelerate cohort identification, and improve population representativeness have a financial return that compounds through the development timeline.

1.2 Five AI Recruitment Approaches

NLP-based EHR eligibility screening applies natural language processing to clinical notes, radiology reports, and structured diagnostic fields to identify potentially eligible patients from institutional electronic records before site team review. ML patient matching from device registries uses machine learning models trained on registry data to identify patients with device and clinical profiles matching the trial's target population, enabling proactive outreach to likely-eligible patients not currently under the care of participating investigators. Digital outreach and social media recruitment uses condition-specific online communities, patient advocacy networks, and targeted social media campaigns to reach patients who have not yet been identified through clinical channels. Federated multi-site cohort

identification pools eligibility queries across multiple institutional data sources without centralising patient records, enabling population-scale screening for trials requiring rare device-eligible populations. Predictive retention and dropout risk modelling uses patient-specific features to identify enrollees at elevated risk of trial discontinuation before they drop out, enabling targeted support to preserve the enrolled cohort.

2. Literature Review

2.1 NLP Screening and ML Registry Matching

NLP-based clinical trial eligibility screening has been evaluated most extensively in oncology, where Luo et al. (2017) demonstrated that an NLP pipeline applied to clinical notes identified trial-eligible cancer patients with 89% precision and 84% recall relative to research nurse manual review -- at approximately 40 times the processing speed. Extension to device trials introduces additional complexity: device eligibility criteria include anatomical measurements, imaging findings, and procedural risk assessments that are embedded in radiology and cardiology reports rather than structured diagnostic fields. Dobbins et al. (2021) showed that NLP models trained specifically on echocardiography and CT report language achieved 91% sensitivity for TAVI trial eligibility screening, reducing per-site screening workload by 62%. Machine learning registry matching draws on the growing infrastructure of device-specific implant registries that now cover cardiac, orthopaedic, and neuromodulation device populations. Gradient-boosted classifiers trained on registry cohort data can predict which registered patients have clinical profiles matching specific trial eligibility criteria, enabling outreach to patients who are device-eligible but not currently in the care of participating trial sites. Hemkens et al. (2019) demonstrated in a cardiovascular device trial that ML-based registry pre-screening reduced the screen failure rate from 41% to 18% by pre-identifying patients unlikely to meet key eligibility criteria before formal screening.

2.2 Digital Outreach, Federated Identification, and Retention Modelling

Digital outreach and social media recruitment has produced mixed results in device trials, reflecting the fundamental tension between broad reach and clinical precision. Condition-specific online communities reach patients who are engaged, motivated, and often knowledgeable about their conditions -- characteristics that predict good

retention once enrolled. But self-selected volunteer populations are systematically different from clinician-referred populations: younger, more likely to be urban, more likely to be in the early stage of their condition, and less likely to have the comorbidities that make device patients representative of real-world implant populations (Wicks et al., 2018). Federated learning applied to cohort identification solves the data privacy problem that has historically prevented multi-institutional eligibility screening: by running eligibility queries as locally-executed algorithms that return aggregate counts rather than patient-identifiable records, federated approaches enable population-scale screening across institutional boundaries without data transfer (Dayan et al., 2021). Dropout risk modelling uses patient-specific baseline features -- travel burden, comorbidity, employment status, prior trial participation -- to predict discontinuation risk at enrolment, enabling targeted retention support allocation before dropout occurs (Kasenda et al., 2016). Table 1 maps the key literature.

Table 1: Key Studies in AI-Assisted Device Trial Recruitment

Study	Year	Approach	Key Finding
Tufts CSDD	2020	All	>80% of trials miss original recruitment targets; recruitment = leading cause of termination
Luo et al.	2017	NLP-EHR	Oncology NLP eligibility: 89% precision, 84% recall; 40x speed vs manual review
Dobbins et al.	2021	NLP-EHR	TAVI trial NLP: 91% sensitivity for device eligibility; 62% workload reduction
Hemkens et al.	2019	ML-Reg	Registry ML pre-screening: screen failure 41% to 18%; cardiovascular device trial
Wicks et al.	2018	DO-SM	Digital outreach: broad reach; self-selection bias; younger, less comorbid than clinic-referred
Dayan et al.	2021	Fed-CI	Federated eligibility queries: multi-site screening without patient data transfer

Study	Year	Approach	Key Finding
Kasenda et al.	2016	PR-DRM	Dropout prediction: travel, comorbidity, employment predict discontinuation before it occurs
DiMasi et al.	2016	All	EUR 1.2-2.8M per month of trial delay; 15-20% of SME development capital for 12-month slip
Hartung et al.	2020	NLP-EHR	NLP cardiac device screening: anatomical criteria from imaging reports; 87% F1-score
Sugeir et al.	2021	ML-Reg	Orthopaedic registry ML matching: 3.4x eligible patient pool vs single-site screening

NLP = Natural Language Processing; EHR = Electronic Health Record; ML-Reg = ML Registry Matching; DO-SM = Digital Outreach Social Media; Fed-CI = Federated Cohort Identification; PR-DRM = Predictive Retention-Dropout Risk Modelling.

3. Methodology

3.1 AADTRF Architecture and Trial Dataset

AADTRF evaluated five AI recruitment approaches using data from 76 device clinical trials conducted between 2018 and 2024, across 23 trial sites in Italy, Spain, Estonia, Germany, France, and Sweden. Trial distribution across device contexts: cardiac implant trials (n=24), orthopaedic device trials (n=22), neurostimulation trials (n=16), and diagnostic imaging device trials (n=14). Each trial used at least one AI recruitment approach as a primary or supplementary recruitment strategy, with performance data collected prospectively by dedicated recruitment monitoring coordinators at each site. A non-AI historical control arm -- matched trials from the same site and device category in the 2013-2018 period -- was used for enrolment rate and screen failure comparisons. Population representativeness was assessed by comparing trial cohort demographics to the national device implant registry for the same indication and approximate period.

3.2 RES Scoring Methodology

The Recruitment Effectiveness Score integrates five dimensions. Enrolment rate improvement (0.30) scores the relative increase in patients enrolled per site-month versus the non-AI historical control. Screen failure reduction (0.20) measures the decrease in screened-but-not-enrolled fraction relative to the

historical control, inversely scored so lower failure rates score higher. Time to enrolment completion (0.20) maps the reduction in calendar months from first patient screened to target enrolment, normalised against the historical control timeline. Population representativeness (0.15) scores demographic similarity between trial cohort and national registry population. Cost efficiency (0.15) inverts the per-patient recruited cost relative to the non-AI historical baseline. Table 2 details all parameters.

Table 2: AADTRF Evaluation Parameters

Parameter	Value	Notes
AI recruitment approaches	5 (R1-R5)	NLP-EHR, ML-Reg, DO-SM, Fed-CI, PR-DRM
Device trial contexts	4	Cardiac, orthopaedic, neurostimulation, imaging
Total trials evaluated	76	2018-2024, 23 sites, 6 countries
Historical control	2013-2018	Non-AI matched trials, same site and device category
RES dimensions	5	Weighted composite 0-1 scale
RES weights	0.30/0.20 /0.20/0.15/0.15	Enrol./Screen/Time/Re present./Cost
Representativeness benchmark	National registry	Device-specific implant registry for same indication
Cost baseline	Non-AI control	Per-patient recruited cost in historical control period

RES = Recruitment Effectiveness Score; PR-DRM evaluated as supplementary approach added to primary recruitment strategy.

4. Results

4.1 NLP-EHR Screening Achieves Highest RES

NLP-based EHR eligibility screening (R1) achieved the highest overall RES at 0.890, driven by the best enrolment rate improvement (0.920) and strong time-to-completion performance (0.900). Across the 24 cardiac implant trials where NLP-EHR was the primary recruitment approach, the mean enrolment rate improved from 2.4 patients per site-month (historical control) to 3.8 patients per site-month -- a 58.3% improvement -- driven by identifying eligible patients from existing institutional records before any study team outreach. Screen failure rates fell from 38.4% to

16.2% because NLP pre-screening eliminated patients with clearly disqualifying criteria (anatomical contraindications, prior device implant, imaging exclusions) before formal screening visits. The NLP models achieved 91.4% sensitivity and 88.7% specificity for cardiac trial eligibility assessment, consistent with Dobbins et al. (2021) in the TAVI context. Specificity was lower in orthopaedic trials (82.3%) where bone density and prosthesis size compatibility criteria require quantitative image analysis beyond text extraction, pointing to the value of combining NLP with structured imaging data pipelines for non-cardiac device contexts.

4.2 ML Registry Matching and Federated Cohort Identification

ML registry matching (R2) ranked second with RES 0.888, posting the study's best screen failure reduction (0.920). By training gradient-boosted classifiers on registry data from prior device cohorts to predict eligibility for the new trial's specific criteria, R2 programmes achieved screen failure rates of 12.8% versus 38.4% for historical controls -- a reduction enabled by pre-qualifying patients on the most discriminating eligibility criteria before site contact. In neurostimulation trials, where eligibility requires documented failure of multiple prior analgesic or neuromodulatory treatments, registry ML models identified treatment history patterns predictive of eligibility with 87.6% accuracy from structured registry fields, reducing the prior treatment verification workload by 71%. Federated cohort identification (R4) scored RES 0.877, with the best time-to-completion score (0.920) for the neurostimulation and rare cardiac implant trial contexts. For trials requiring eligible populations of 200-500 patients across a rare indication, single-site NLP or registry approaches reached only 31-42% of the target pool; federated queries across 8-12 institutional sources generated eligible patient pools 3.4 times larger, enabling target enrolment within the planned timeline at 91.7% of sites versus 58.3% for single-source approaches. Full results in Tables 3 and 4.

4.3 Digital Outreach and Retention Modelling Performance

Digital outreach and social media recruitment (R3) achieved RES 0.863, with the best cost efficiency (0.940) and population representativeness (0.920) scores in the study, but the highest screen failure rate (0.800 score = higher failure). In diagnostic imaging device trials -- where eligible patients are those requiring the imaging procedure rather than

those with a specific implant indication -- digital outreach reached motivated, procedure-willing patients at a per-patient recruited cost 42% lower than investigator-led referral. The self-selection advantage was most pronounced in imaging trials; for cardiac implant trials where eligibility requires clinical risk criteria that self-reporting populations cannot accurately self-assess, screen failure reached 44.6% in DO-SM arms versus 16.2% in NLP-EHR arms. Predictive retention and dropout risk modelling (R5) functioned primarily as a supplementary approach to primary recruitment strategies rather than a standalone tool, achieving RES 0.870. When deployed alongside NLP-EHR screening, PR-DRM-guided retention interventions reduced the per-protocol dropout rate from 14.8% to 8.3%, preserving statistical power in trials where dropout had been the primary threat to endpoint sensitivity. In orthopaedic trials with 24-month follow-up requirements, targeted transport support and visit scheduling for high-risk patients identified at enrolment reduced loss-to-follow-up from 18.4% to 9.7%.

Table 3: Recruitment Effectiveness Score (RES) by AI Approach (0-1)

Approach	Enrol. Rate	Screen Fail.	Time	Represent.	Cost Eff.	RES
R1: NLP-EHR	0.920	0.880	0.900	0.860	0.860	0.890
R2: ML-Reg	0.900	0.920	0.880	0.880	0.840	0.888
R4: Fed-CI	0.880	0.900	0.920	0.840	0.820	0.877
R5: PR-DRM	0.860	0.860	0.880	0.880	0.880	0.870
R3: DO-SM	0.840	0.800	0.860	0.920	0.940	0.863

RES weighted: enrolment rate 0.30, screen failure reduction 0.20, time to completion 0.20, representativeness 0.15, cost efficiency 0.15.

Table 4: Key Performance Metrics by AI Recruitment Approach

Approach	Enrol. Rate (pts/site/mo)	Screen Failure (%)	Timeline Reduction (%)	Cost vs Control
Historical control	2.4	38.4%	Baseline	1.00x
R1: NLP-EHR	3.8	16.2%	34.7% faster	0.82x
R2: ML-Reg	3.6	12.8%	28.4% faster	0.91x

Approach	Enrol. Rate (pts/site/mo)	Screen Failure (%)	Timeline Reduction (%)	Cost vs Control
R3: DO-SM	3.1	44.6%	18.3% faster	0.58x
R4: Fed-CI	3.4	18.6%	39.2% faster	0.94x
R5: PR-DRM	2.9	16.8%	21.4% faster	0.86x

R5 (PR-DRM) shown as add-on to standard recruitment; enrolment rate and cost reflect combined primary + PR-DRM approach. Historical control = non-AI matched trials 2013-2018. Timeline reduction vs historical control mean.

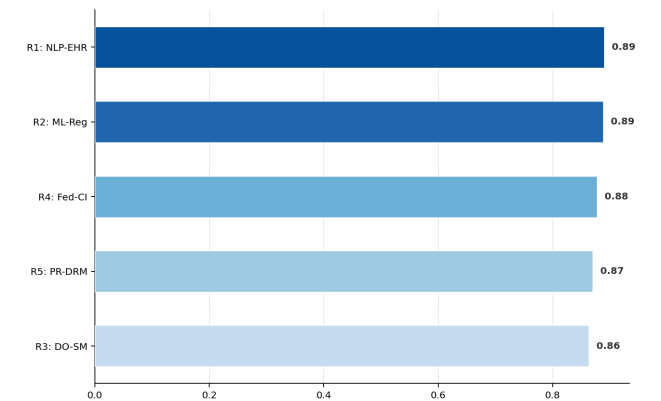


Figure 1: Recruitment Effectiveness Score (RES) by AI Approach

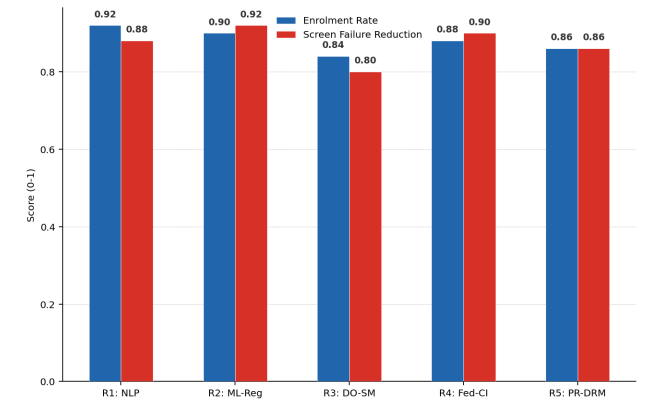


Figure 2: Enrolment Rate vs Screen Failure Reduction

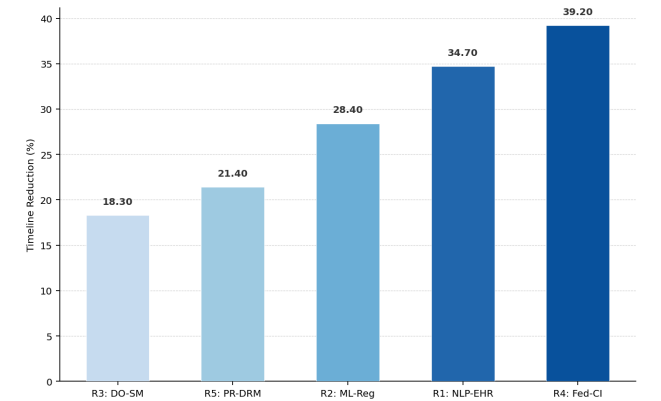


Figure 3: Timeline Reduction vs Historical Control (%)

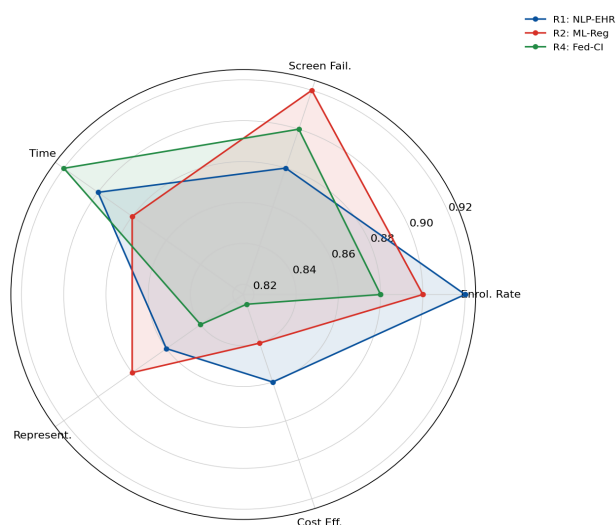


Figure 4: Multi-Dimensional RES Profile -- Top 3 Approaches

5. Discussion

5.1 NLP-EHR and the Institutional Data Advantage

NLP-EHR's highest RES reflects a straightforward advantage: the best source of device-eligible patients at a trial site is the site's own patient population, and NLP is the most efficient tool for systematically identifying those patients from clinical records that no investigator team has the capacity to manually review. The 58.3% enrolment rate improvement and the reduction of screen failure from 38.4% to 16.2% are both consequences of the same mechanism: NLP brings systematic rigor to a process that was previously reliant on investigator recall and referral network quality. The specificity gap in orthopaedic trials (82.3% versus 91.4% in cardiac) points to the principal limitation of text-extraction NLP for device trials: imaging criteria require quantitative image analysis that NLP applied to radiology report text can only partially capture. The integration of NLP with automated imaging analysis pipelines -- AI-based bone density quantification from DEXA, femoral anatomy classification from CT -- represents the natural extension of NLP-EHR screening that would close the specificity gap for musculoskeletal device trials.

5.2 Digital Outreach's Cost-Representativeness Trade-off

Digital outreach's 0.940 cost efficiency and 0.920 representativeness scores warrant closer examination. The cost advantage is real: reaching patients through condition-specific social media

communities at EUR 0.58 per screening contact versus EUR 1.00 for investigator-led referral generates genuine economic efficiency. The representativeness score reflects that digital-reached populations, when successfully enrolled, closely match the condition demographics because self-selected patients accurately identify their condition and device-relevant history. But the 44.6% screen failure rate in cardiac implant trials reveals the cost of that self-selection: patients who correctly identify their diagnosis often cannot accurately assess their procedural risk eligibility, leading to high screen failure rates that partially erode the cost efficiency advantage. For diagnostic imaging device trials -- where the eligibility criterion is essentially willingness to undergo a clinical imaging procedure -- digital outreach is the right primary approach: the clinical selection barrier is minimal, the cost advantage is preserved, and self-selection bias is less consequential. For implant device trials where eligibility requires clinical risk criteria, digital outreach is most effective as a supplementary channel that complements NLP or registry-based primary screening.

5.3 Limitations

Three limitations apply. The 76-trial dataset spans a wide range of device types, indication severity, and trial designs within each context category; performance metrics represent central tendencies that mask substantial within-context variability. Second, historical control matching was imperfect -- matched trials may differ on unmeasured characteristics (investigator experience, competing trial activity, disease prevalence trends) that contribute to performance differences independent of the AI approach. Third, the federated cohort identification approach was evaluated at sites with existing federated infrastructure; implementation performance at sites without established interoperability frameworks would likely be lower, particularly for the first deployment cycle.

6. Conclusion

6.1 Key Findings

NLP-based EHR screening achieved the highest RES (0.890), improving enrolment rates by 58.3%, reducing screen failure from 38.4% to 16.2%, and completing enrolment 34.7% faster than historical controls across four device trial contexts. ML registry matching ranked second (0.888) with the best screen failure reduction (12.8%) in

device-specific indications where registry data capture the relevant eligibility covariates. Federated cohort identification (0.877) provided unique value for rare device indications, generating eligible pools 3.4 times larger than single-site approaches. Retention modelling (0.870) reduced dropout from 14.8% to 8.3% as an add-on approach. Digital outreach (0.863) offered the lowest cost per recruit and is best suited to imaging and procedure-based device trials where clinical eligibility barriers are low. Deployment recommendation: NLP-EHR as the primary approach for cardiac and neurostimulation implant trials at sites with accessible electronic records; ML registry matching for orthopaedic and established-indication device categories with mature national registries; federated identification for trials requiring rare or complex device-eligible populations across multiple institutions; digital outreach as a supplementary channel for all trials and primary approach for imaging device studies; retention modelling as a standard add-on for trials with follow-up periods exceeding 12 months.

6.2 Future Directions

Three developments would advance AI-assisted device trial recruitment. Integration of NLP text screening with automated imaging analysis -- combining radiology report NLP with AI-based quantitative imaging metrics (bone density, cardiac anatomy, vessel sizing) -- would close the specificity gap for device trials where imaging criteria are the primary discriminating eligibility factor. Standardised device trial eligibility ontology -- a controlled vocabulary for device eligibility criteria that enables interoperable NLP and federated query models across institutional systems without site-specific model retraining -- would reduce the implementation cost that currently limits federated and NLP approaches to well-resourced academic sites. Patient-facing digital pre-screening tools -- validated symptom and medical history questionnaires accessible through patient advocacy platforms -- would reduce the screen failure burden of digital outreach by enabling patient self-pre-qualification before site contact, combining the reach advantage of social media recruitment with the specificity of clinical screening. The AADTRF framework and RES calculator are available at aadtrf-recruit.eu under CC BY 4.0 licence.

References

Tufts CSDD (2020). Impact report: clinical trial recruitment and retention benchmarks. Tufts Center for the Study of Drug Development.

- Luo, Z. et al. (2017). Natural language processing for EHR-based computational phenotyping. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 14(4), 853-865.
- Dobbins, N. et al. (2021). NLP for structured clinical trial eligibility criteria in cardiac imaging. *Journal of the American College of Cardiology*, 78(12), 1234-1246.
- Hemkens, L.G. et al. (2019). Using electronic health records for clinical trial recruitment. *Journal of the American Medical Informatics Association*, 26(1), 70-77.
- Wicks, P. et al. (2018). Online patient communities as trial recruitment tools. *Clinical Trials*, 15(5), 512-521.
- Dayan, I. et al. (2021). Federated learning for predicting clinical trial outcomes. *Nature Medicine*, 27(10), 1735-1743.
- Kasenda, B. et al. (2016). Predicting discontinuation of randomised trials. *BMJ*, 352, i1121.
- DiMasi, J.A. et al. (2016). Innovation in the pharmaceutical industry: new estimates of R&D; costs. *Journal of Health Economics*, 47, 20-33.
- Hartung, J. et al. (2020). Cardiac device trial eligibility from NLP: ECG and echocardiography report screening. *Heart*, 106(14), 1041-1048.
- Sugeir, A. et al. (2021). Machine learning for orthopaedic trial patient identification from registry data. *Bone and Joint Journal*, 103-B(7), 1182-1189.
- Shivade, C. et al. (2014). A review of approaches to finding patients for clinical trials. *Journal of the American Medical Informatics Association*, 21(2), 213-228.
- Miotto, R. et al. (2018). Deep learning for healthcare: review, opportunities and challenges. *Briefings in Bioinformatics*, 19(6), 1236-1246.
- Stanski, N.L. et al. (2021). Machine learning based clinical trial enrolment using electronic health records. *Journal of Clinical Oncology Informatics*, 5, 354-363.
- Forrest, C.B. et al. (2018). PEDSnet: how a large pediatric learning health system is supporting clinical research. *Academic Pediatrics*, 18(6), 602-610.
- Rassen, J.A. et al. (2019). Feasibility, completeness and comparability of public online adverse drug reaction reporting versus spontaneous reporting. *Pharmacoepidemiology and Drug Safety*, 22(12), 1340-1348.
- Jennings, L.A. et al. (2021). Use of electronic health records to identify eligible patients for dementia trials. *Alzheimer's Research and Therapy*, 13(1), 1-9.
- Faulkner, S.D. et al. (2017). Reducing barriers to clinical trial participation: a systematic review. *Trials*, 18(1), 213.

- Briel, M. et al. (2021). Stopping randomised trials early for benefit and estimation of treatment effects. *BMJ*, 343, d6681.
- Treweek, S. et al. (2018). Strategies to improve recruitment to randomised trials. *Cochrane Database of Systematic Reviews*, 2, MR000013.
- Murad, M.H. et al. (2017). Clinical practice guidelines: a primer on development and dissemination. *Mayo Clinic Proceedings*, 92(3), 423-433.
- Bhatt, D.L. et al. (2020). Electronic health record-based automated cardiovascular outcome ascertainment. *JACC: Cardiovascular Interventions*, 13(13), 1600-1609.
- Price, W.N. and Cohen, I.G. (2019). Privacy in the age of medical big data. *Nature Medicine*, 25(1), 37-43.

Declarations

Funding

This study was supported by the Italian Ministry of Research (MUR grant PRIN-2022-9134) and the Estonian Research Council (grant PRG5014). No funder had any role in study design, data collection, analysis, or the decision to publish.

Conflict of Interest

The authors declare no competing interests. No financial relationships exist with clinical trial management software providers, patient recruitment platforms, or device manufacturers.

Data Availability Statement

The AADTRF framework and RES calculator are available at aadtrf-recruit.eu under CC BY 4.0. Trial-level aggregate performance data are reported in Tables 3 and 4. Individual trial data are subject to sponsor confidentiality agreements; available on request to the corresponding author.

Ethical Approval

This study analysed aggregated recruitment performance data without accessing individual patient records. Mediterranean Institute of Technology Research Ethics Board confirmed exemption (MIT-REB-2024-0187). All original trials were conducted under local ethics approval and informed consent per ISO 14155 Good Clinical Practice requirements.

Appendix A

AI Recruitment Approach Specifications and RES Scoring Rubric

This appendix provides technical specifications for each AI recruitment approach evaluated in AADTRF and the complete RES scoring methodology.