European Journal of
Medical Research and Clinical Trials

JScholar
Fostering Scholarly Communication

Research Article Open Access

A Novel Clinical Trials Search Tool with Real-time Iterative and Geospatial Capabilities

George Udeani[1*], Joy Alonzo[1], Keith Biggers[1], Sergio Herrera[1], Chioma Ogbodo[1], Dorothy Farrell[2], Ladan Panahi[1], Miranda Steinkopf[2], Matthew Barry[1], Terri Moore[2] and Marcia G Ory[1]

[1]Clinical Professor, Department of Translational Medical Sciences and Center for Genomic & Precision Medicine, School of Medicine, Texas A&M University, Texas, United States of America

[2]Clinical Professor, Department of Pharmacy Practice, Irma Lerma Rangel School of Pharmacy, American Association of Colleges of Pharmacy, United States of America

[*]**Corresponding Author:** George Udeani, Clinical Professor, Department of Translational Medical Sciences and Center for Genomic & Precision Medicine, School of Medicine, Texas A&M University, Texas, United States of America, Tel: +19795746440, E-mail:udeani@tamu.edu, chioma_1991@tamu.edu

Received Date: August 01, 2025 **Accepted Date:** August 17, 2025 **Published Date:** August 20, 2025

Citation: George Udeani, Joy Alonzo, Keith Biggers, Sergio Herrera, Chioma Ogbodo, et al. (2025) A Novel Clinical Trials Search Tool with Real-time Iterative and Geospatial Capabilities. Eur J Med Res Clin Trials 7: 1-10

Abstract

Objective: To develop a web-based custom querying and reporting Clinical Trials Search Tool (CTST), built upon ClinicalTrials.gov data, with interactive and geospatial mapping capabilities.

Methods: The CTST was developed using a rapid, iterative, feedback-focused Agile-based software engineering methodology. The tool is broken down into two components. The first is a front-end web client that enables specification of search details and renders results. The second is a back-end server that handles the searches and returns the associated results in a ranked fashion. The front-end web client is written in Svelte and TypeScript using the SvelteKit framework and Tailwind CSS User Interface (UI) library. It communicates with the server's REST API using asynchronous HTTP requests containing JSON-serialized data. The back-end server is written in Java using the Spring Boot framework and stores application data in a PostgreSQL database.

Results: The CTST provides an easy-to-use interface for specifying input search criteria. This allows the user to select only the criteria that will best match their areas of interest. Clicking the "Toggle Advanced Search" button leads to the keyword search box revealing additional search criteria for filtering search results. An interactive geospatial map renders the returned

results as listings that summarize clinical trials in a tabular form, several layers that integrates population density, racial and ethnicity demographics, and socioeconomic details.

Conclusion: The CTST builds upon the ClinicalsTrials.gov database, leveraging its geospatial and other capabilities for advanced research purposes, thus optimizing its use by clinicians and other investigators.

Keywords: Clinical Trials; Study; Tool; ClinicalTrials.gov; Geospatial; Iterative; Portal; Location; Dermographics; Condition

Introduction

Each year, several drugs are approved by the United States Food and Drug Administration (FDA) for clinical use. In 2024, the Center for Drug Evaluation and Research (CDER) approved 50 novel drugs [1]. The journey of the tens of thousands of drug candidates that go through the drug development process for ultimate approval begins at several university laboratories, teaching hospitals, private entities, globally, and US government agencies, culminating in various phases of clinical trials.

In 2024, the US clinical trials market was valued at USD 40.56 billion and is expected to grow at a compound annual growth rate (CAGR) of 6.30 % between 2025 and 2030. This market growth is fueled by the increasing incidence of chronic ailments and the demand for clinical trials in emerging economies. Other factors include the growing number of biologics, precision medicines, orphan drugs, and the demand for advanced technologies [2].

The Advanced Research Project Agency for Health (ARPA-H) is a research financing agency that supports transformative biomedical and health breakthroughs, from molecular to societal, to provide health solutions for all. This agency promotes the nation's ability to conduct clinical trials equitably and enhance clinical trial access for people in communities across the United States. Advancing Clinical Trial Readiness (ACT-R) aims to make clinical trials accessible within 30 minutes to 90% of eligible Americans [3,4]. ARPA-H has called for this to be driven through technology, such as artificial intelligence and machine learning. For this to be achieved, ACT-R has developed the ARPANET-H Consumer Experience Hub. ARPA-H intends to harness clinical trial potentials that surmount challenges in assessing new technologies, therapies, and platforms through projects with different collaborators. Technological advancements such as geospatial intelligence and machine learning are geared towards optimizing clinical trial recruitment infrastructure. At Texas A&M University, we have previously employed artificial and geospatial intelligence to determine the proximities of COVID-19 clinical trials and pharmacies to US populations [5].

The ClinicalsTrials.gov website provides a valuable resource for obtaining information about ongoing or concluded clinical trials [6,7] This site provides basic tools for searching and obtaining the associated details around clinical trials of interest but does not include other supporting data needed to properly query the platform to understand the surrounding details, such as real-time population types and densities as well as proximity data. At Texas A&M University, we developed a Clinical Trials Search Tool (CTST) that uses data from ClinicalsTrials.gov to provide a more comprehensive index and an easy-to-use resource for clinicians, researchers, and other interested parties. This tool functions as a web-based application that enables custom querying trials based upon different search criteria, returning results ranked based on their relevance to the specified criteria and displaying the returned results in a geospatial and tabular summary fashion. It also integrates other supporting geospatial data of interest, including demographic and socioeconomic details. This tool aims to provide users with a quick and easy method to find clinical trials of interest while leveraging details around them. Given the ARPA-H intent to have clinical trials within 30 minutes of 90% of eligible Americans, this tool can geospatially compute the proximity of clinical trials from 20 to 100 miles of potential clinical trial populations.

The modernized ClinicalTrials.gov site delivers a streamlined search experience, enhanced filters, and a map-

based location search, improving how users find studies near a city, address, or current location. While these updates elevate usability, ClinicalTrials.gov remains focused on study registry data and does not natively incorporate census-level demographic or socioeconomic overlays into its search interface [8] By contrast, CTST augments the public registry with investigator-oriented geospatial layers (e.g., population density, race/ethnicity, income) to support feasibility assessment and site strategy alongside conventional keyword and filter queries. This approach complements, rather than duplicates, recent ClinicalTrials.gov modernization efforts.

Patient-facing recruitment platforms such as Antidote and TrialX emphasize matching patients to trials and sponsor workflows (e.g., guided search, pre-screeners, enrollment support), rather than investigator geospatial enrichment [9]. Carebox likewise focuses on connecting patients, families, and physicians to relevant studies with technology for navigation and matching [10]. Enterprise networks like TriNetX provide cohort discovery and site selection by querying federated EHR data within a member ecosystem, enabling de-identified cohort counts that local sites can re-identify for outreach [11]. In contrast, CTST ingests and cross-references data publicly available from ClinicalTrials.gov and the US Census Bureau. The combination of these public datasets provides geospatial overlays, flexible search utilities backed by Elasticsearch, and feasibility-oriented mapping without requiring participation in a proprietary data network. CTST does not keep its own copies of these datasets, but rather indexes them and provides links to the original data where applicable. The CTST requires authentication via Google's identity provider using OIDC (OpenID Connect) / OAuth 2.0 due to the tool's potential to house sensitive data in the future. Protected data could include: i) user account and institutional affiliation details, ii) saved searches, custom site lists, and investigator notes, iii) sponsor-provided feasibility metrics and site performance data, and iv) optional EHR-linked eligibility outputs that remain de-identified at the CTST layer. These data are protected because they may contain PII (personally identifiable information)/PHI (protected health information), proprietary sponsor information, or, when combined with high-resolution geospatial layers, could increase re-identification risk. The CTST provides a centralized resource that manages several aspects of clinical trials, including the dates of the clinical trial, participants' demographics, study type (interventional, observational, and expanded access), disease states, status (completed, unknown, recruiting, terminated), geospatial results (distances, population density, race, ethnicity, and socioeconomic layers).

Methods

The CTST was developed using an Agile-based software engineering methodology [12], which involves a rapid, iterative, feedback-focused software generation. The result is a finished product that addresses the requirements and feedback from subject-matter experts to the end-user. This process enables iterative development and refinement of the tool and allows frequent feedback collection on an incremental and expanded set of features and capabilities.

The tool is broken down into two components: i) a front-end web client that enables specification of search details and rendering of returned results, and ii) a back-end server that handles the searches and returns the associated results in a ranked fashion. The front-end web client is written in Svelte and TypeScript using the SvelteKit framework and Tailwind CSS User Interface (UI) library [13]. It communicates with the server's REST API using asynchronous HTTP requests containing JSON-serialized data [14] The back-end server is written in Java using the Spring Boot framework and stores application data in a PostgreSQL database.

The back-end server builds on an Elasticsearch index for efficient searching and retrevial of relevant records within the Clinical Trials dataset. The Clinical Trials dataset is automatically downloaded from the ClinicalsTrials.gov website periodically, unpacked, and then processed into the Elasticsearch index. This index consists of documents in the form of key-value pairs (i.e., stored as JSON documents), and this index can be searched through a RESTful JSON-based API. A set of ranked records (i.e., scored and ordered based on their relation to the input search criteria) are returned as JSON, and these records are then passed on to the front-end web client for display. To conserve space and optimize efficiency, the server only stores and transmits the minimum data required to display a list of search results. Users

are referred to the canonical upstream sources for complete details (i.e., linking to the original record for a selected trial in ClinicalsTrials.gov website for the full details). A high-level architecture is provided in Figure 1.

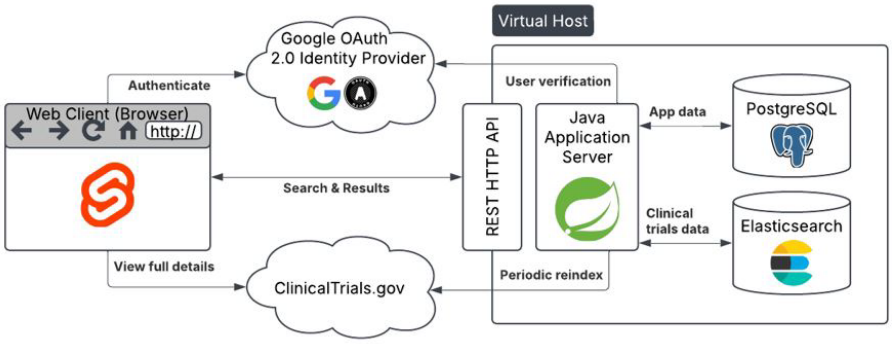

Figure 1: CTST Search High-Level Architecture

The front-end web client allows for the definition of searches in one of two modes: i) a basic search that uses basic keywords and ii) an advanced search that incorporates different faceted criteria. Search criteria include:

- **Keywords:** terms that can be matched against the title and detailed description for the clinical trial
- **Date Range:** dates in which the clinical trial is enrolling participants
- **Location:** location(s) where the clinical trial is enrolling participants
- **Participant Demographics:** racial and ethnic makeup of the enrolled participants
- **Study Type:** the type of trial (i.e., interventional, observational, expanded access)
- **Conditions:** any specialized conditions being addressed
- **Status:** active status of the clinical trial (e.g., not yet recruiting, recruiting, completed, terminated)

CTST internally uses Elasticsearch, which in turn is built on Lucene, to identify trials of interest to the user. It ranks trials using a variety of text, numeric, and geographic indexing approaches across key fields (title, conditions, brief summary, interventions, MeSH). Relevance is then calculated via similarity to the user's query taking into account term frequency, inverse document frequency, and field--length normalization with optional faceted filters for incorporating additional fields.

Once the back-end server has executed the search, results are returned to the front end and displayed in several ways. This includes a geospatial depiction of the results as well as a tabular summary with the key details highlighted (i.e., the title, sponsor, status, phase, type, and conditions for each trial, along with a reference link to a trial's full details on the ClinicalsTrials.gov website). Multiple deployment modes were developed for the tool. One version allows users to search only records related to COVID-19, while another provides access to the entire ClinicalTrials.gov dataset. Additional future versions can be readily configured to target searches on specific topics or areas of interest. The COVID-19 version of the tool is publicly accessible; however, users must authenticate to access the full dataset search tool. Authentication is handled through Google's OAuth 2.0 API, which verifies the user's identity. Upon first login, access is confirmed by the server using an invitation code.

The CTST features an intuitive interface for entering search criteria (Figure 2). A basic search functions similarly to standard online search tools, matching the entered terms with clinical trial details and returning ranked results based on relevance. The advanced search offers users a more powerful and flexible way to define and refine their search queries. It is important to note that in an advanced search, users are not required to provide values for every cri-

terion; however, all criteria that are specified will be taken into account during the search. Users can select only the criteria that align with their specific interests—for example, retrieving trials conducted within a particular timeframe, at a specific location, or based on a specific study type or targeted conditions. Clicking the "Toggle Advanced Search" button beneath the keyword search box reveals additional filtering options. The keyword field can be left blank if the user chooses to search using only these advanced criteria. After entering the desired parameters, clicking the blue magnifying glass icon to the right of the search box will initiate the search.

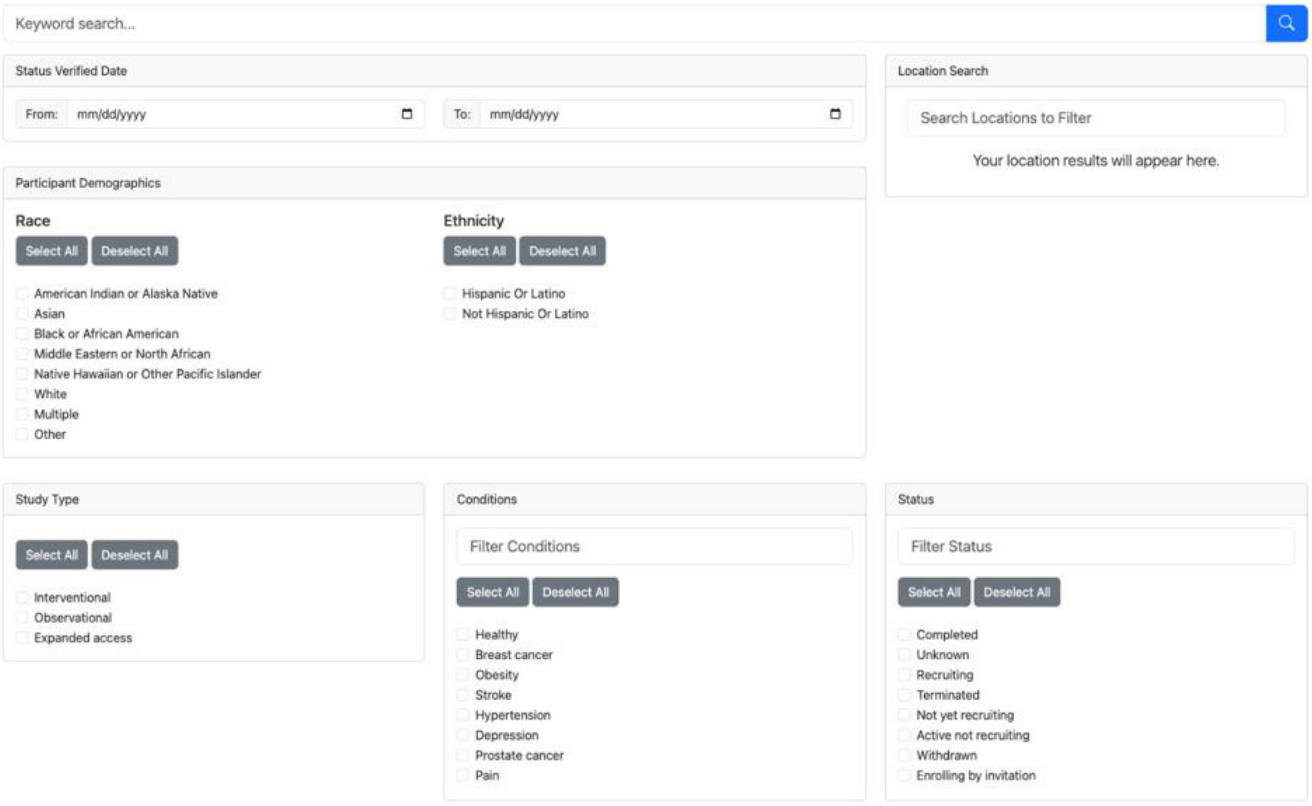

Figure 2: CT Search Criteria

After the search is successfully executed and the results are returned to the front end, they are presented in multiple formats. First, an interactive geospatial map displays the results (Figure 3). The orange pins on the map indicate the locations where matching clinical trials are being conducted, with the numeric value representing the number of trials in each area. It is important to note that a single trial may span multiple locations, and some trials may lack location data altogether. The map is fully interactive, and zooming in will cause clustered pins to separate, revealing more detailed information for each specific area. The map also features multiple toggleable layers that enhance its functionality. These layers incorporate data from the U.S. Census Bureau, including population density, racial and ethnic demographics, and socioeconomic indicators such as average and median household income, median age, and gender distribution. Users can activate one or more layers simultaneously and adjust their opacity to blend them visually, enabling cross-referencing of clinical trial locations with relevant demographic and socioeconomic data. Layer controls and opacity sliders are located on the left side of the map, allowing users to modify the visibility and color intensity of each layer. Clicking on a pin or a county reveals detailed information about the selected item in the panel on the right. Additionally, selecting a clinical trial name in this panel highlights the corresponding trial in the results listing section below the map.

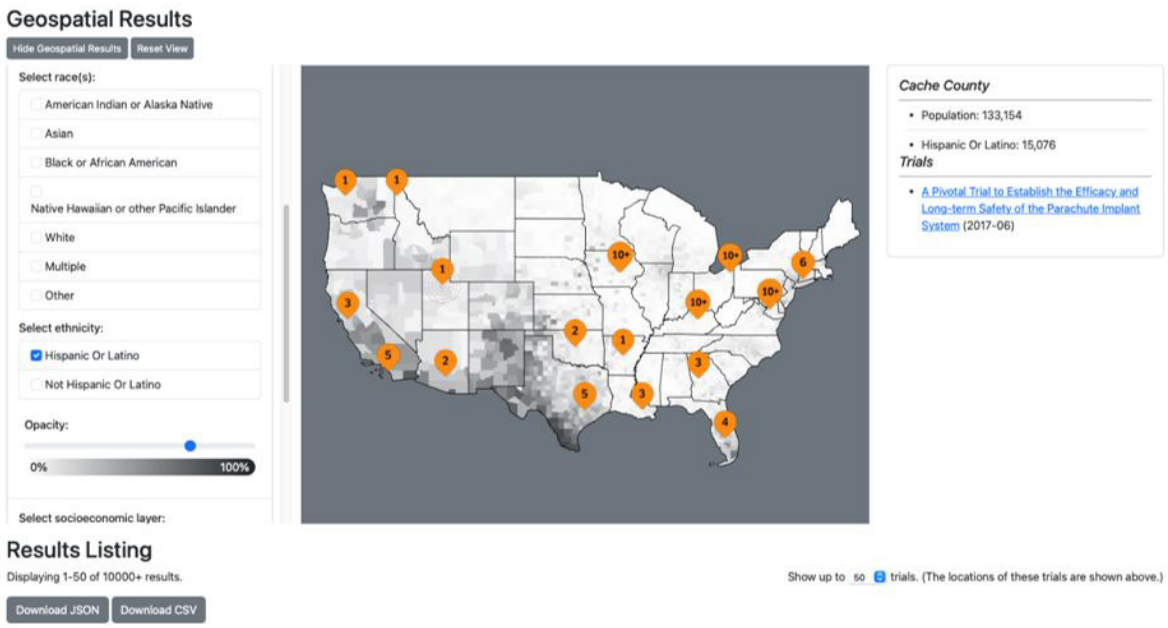

Figure 3: Search Results in a Geospatial Fashion

In addition to the map view, search results are presented in a "Results Listing" format, where each clinical trial is summarized in a tabular layout (Figure 4). Trials are ordered by a "match score", a relative value where higher scores indicate a stronger relevance to the search criteria. Results are ordered by a determined match score, where higher values indicate closer alignment between the query and trial record; details of the scoring method are provided in Methods. To optimize performance and avoid overloading the application, the listing displays a limited number of results per page (10 by default). Users can adjust the number of results shown per page using the drop-down menu located to the right of the Results Listing header.

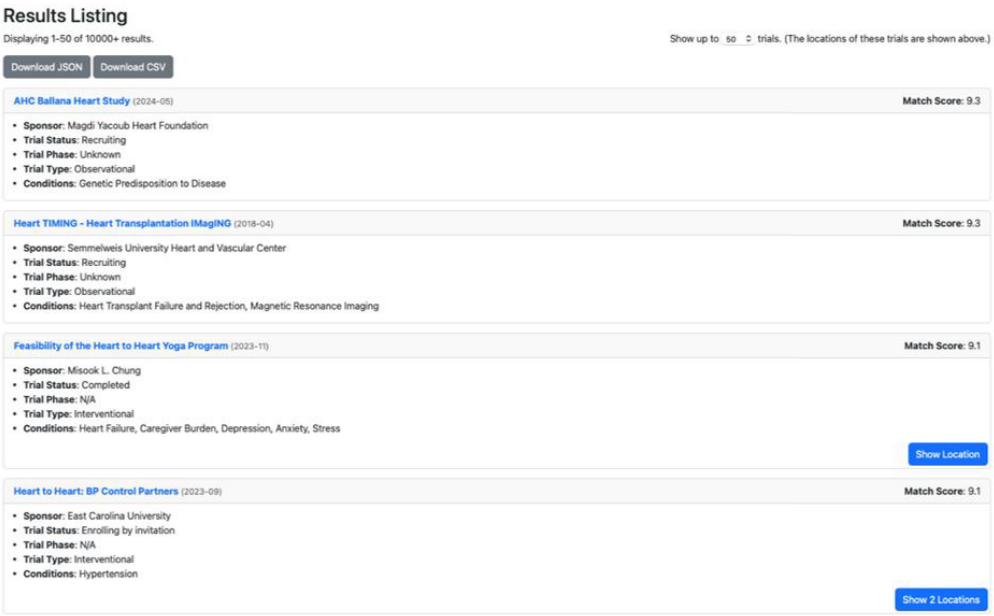

Figure 4: Search ResultsDisplayed in a Tabular Summary Format

The Results Listing section presents matching clinical trials in a detailed list format, with specific information provided for each trial. For trials that include location data, a map is available showing all sites where the trial is being conducted. Clicking the "Show Location(s)" button, when present, will display this map (Figure 5).

Figure 5: Individual Clinical Trial Result

Discussion

Clinical trials are an integral part of drug discovery and development [15] The US clinical trials market was valued at USD 40.56 billion in 2024 and is expected to grow at a compound annual growth rate (CAGR) of 6.30 % between 2025 and 2030.2 Clinicaltrials.gov is a clinical trials database published and maintained by the National Library Of Medicine, which houses ongoing and completed clinical trials in the US, and internationally [16]. There are two major reasons for clinical trial registration and results submission to ClinicalTrials.gov. The first is that policies and laws mandate certain clinical trial registration [17]. These include: i) Final Rule (42 CFR Part 11) for the Food and Drug Administration Amendment Acts of 2007 (FDAAA), ii) NIH Policy on Dissemination of NIH-funded Clinical Trial Information, and iii) WHO International Clinical Trials Registry Platform are published on ClinicalTrials.gov. The second reason is the International Committee of Medical Journal Editors (ICMJE) Policy requirement for publishing results from clinical trials in medical journals [18,19]. Under the FDAA, basic results of clinical trials, which include drugs or devices that the FDA has approved are reported on ClinicalTrials.gov website [20].The ClinicalTrials.gov database was also developed to enhance public access to clinical trials. The platform provides information about ongoing and completed studies and their results to researchers, healthcare professionals, and the public [21]. Also, the timely publication of clinical trial results is necessary for increasing transparency and guaranteeing that the research findings are available to inform evidence-based medicine [22,23]. In 2024, the 500,000th clinical study was recorded on ClinicalTrials.gov [24], While ClinicalTrials.gov has been a valuable resource for about 25 years, it has some limitations, which include a lack of integration of clinical trials proximity, as well as robust demographic and socioeconomic data. Given ACT-R's intent, to make clinical trials accessible within 30 minutes to 90% of eligible Americans,3 the CTST could be leveraged to advance this objective. The CTST is built upon the ClinicalTrials.gov database and enhances it by providing detailed geospatial insights. Unlike ClinicalTrials.gov, this web-based platform integrates trial location, geographic proximity, population demographics, and socioeconomic data, offering investigators a more comprehensive view of the clinical trial landscape. The tool enables rapid access to relevant trials and supports efficient patient identification and enrollment based on specific research interests. Users can tailor their searches by selecting criteria aligned with their study focus. The tool then generates a list of ongoing or completed trials, including informa-

tion on location, targeted populations, and the disease conditions under investigation. Search results are displayed geospatially via an interactive map showing trial density by region and in a sortable, tabular format. Each listing includes a match score, where higher scores indicate more substantial alignment with the user's search parameters. Zooming in on the map reveals more granular details about specific trial locations.

Conclusion

The CTST enhances the ClinicalTrials.gov database by incorporating geospatial intelligence, including location data, geographic proximity, population demographics, and socioeconomic indicators. This integration provides investigators with a more comprehensive dataset for trial planning, patient identification, and enrollment aligned with specific research interests. In the context of the rapidly expanding clinical trials landscape and in alignment with ACT-R's goal of making clinical trials accessible within 30 minutes for 90% of eligible Americans, the CTST offers a valuable resource for generating robust, real-time data to inform and accelerate clinical research. To further strengthen these contributions, we recommend extending CTST with mobile access for point-of-care and community use, integration with EHRs for automated eligibility screening and cohort discovery, multilingual support to engage diverse populations, and AI-powered recommendations to optimize site selection, outreach strategies, and patient-trial matching.

References

1. United States Food And Drug Administration (2025) Novel Drug Approvals for 2024. FDA.

2. Grand View Research I (2025) U.S. Clinical Trials Market Size, Share & Trends Analysis Report By Phase (Phase I, Phase II, Phase III, Phase IV), By Indication, By Study Design, By Indication by Study Design, By Sponsor, By Service, And Segment Forecasts, 2025 - 2030. U.S. Clinical Trials Market Size.

3. Advanced Research Projects Agency for Health (2025) ARPA-H advances initiative to improve clinical trials. ARPA-H.

4. Collins FS, Schwetz TA, Tabak LA, Lander ES (2021) ARPA-H: Accelerating biomedical breakthroughs. Science. 373: 165-7.

5. Ogbodo C, Alonzo J, Moore T, Ory M, Udeani G (2025) Pharmacy Advances Clinical Trials (PACT) Network to Achieve Diversity in COVID Clinical Trials: A Strategic Framework. Original Research- Conference Abstract. American Journal of Health-System Pharmacy. 82: S1521.

6. Zarin DA, Tse T, Williams RJ, Califf RM, Ide NC (2011) The ClinicalTrials. gov results database—update and key issues. New England Journal of Medicine. 364: 852-60.

7. National Library of Medicine. ClinicalTrials.gov is a place to learn about clinical studies from around the world. National Library of Medicine.

8. National Library of Medicine (2025) Report on the ClinicalTrials.gov modernization effort.

9. TrialX (2025) Enhance your clinical trial patient recruitment efforts with TrialX clinical trial finders.

10. Health C (2025) Carebox—connect patients, families, and physicians to clinical trials.

11. TriNetX (2025) Real-world data for the life sciences and healthcare—platform overview.

12. Chan FK, Thong JY (2009) Acceptance of agile methodologies: A critical review and conceptual framework. Decision support systems. 46: 803-14.

13. Klimm MC (2021) Design Systems for Micro Frontends-An Investigation into the Development of Framework-Agnostic Design Systems using Svelte and Tailwind CSS. Hochschulbibliothek der Technischen Hochschule Köln.

14. Koubaa A (2015) ROS as a service: web services for robot operating system. Journal of Software Engineering for Robotics. 6: 1-14.

15. Piantadosi S (2024) Clinical trials: a methodologic perspective. John Wiley & Sons.

16. Boston University Medical Campus and Boston Medical Center (2025) Office of Human Research Affairs. ClinicalTrials.gov – What, Why, Which Studies, When. Boston University Medical Campus and Boston Medical Center:Office of Human Research Affairs.

17. Zarin DA, Fain KM, Dobbins HD, Tse T, Williams RJ (2019) Ten-year update on ClinicalTrials. gov results database. The New England journal of medicine. 381: 1966.

18. National Library of Medicine (2025) Clinical Trial Reporting Requirements. National Library Of Medicine.

19. Zarin DA, Tse T, Williams RJ, Carr S (2016) Trial reporting in ClinicalTrials. gov—the final rule. New England Journal of Medicine. 375: 1998-2004.

20. Tse T, Williams RJ, Zarin DA (2009) Reporting "basic results" in ClinicalTrials. gov. Chest. 136: 295-303.

21. National Library of Medicine (2025) About ClinicalTrials.gov. National Library Of Medicine.

22. Nelson JT, Tse T, Puplampu-Dove Y, Golfinopoulos E, Zarin DA (2023) Comparison of availability of trial results in ClinicalTrials. gov and PubMed by data source and funder type. JAMA. 329: 1404-6.

23. Lynch HF, Largent EA, Zarin DA (2017) Reaping the bounty of publicly available clinical trial consent forms. IRB. 39: 10.

24. National Library of Medicine (2025) ClinicalTrials.-

gov: A 25-Year Journey to a Half-Million Registered Studies NIH.