

Clinical Trials Browser: Exploring the ClinicalTrials.gov Registry

Colin Gerber, Anobel Odisho, Jason Ost

Background

Clinical studies include interventional and observational studies of human subjects in order to increase medical knowledge.¹ Interventional studies, also known as clinical trials, test the effect of a specific drug, device, procedure, or behavior change, typically in comparison to some known treatment or a placebo (i.e., no treatment or intervention). Clinical trials are conducted in a series of phases:²

- Phase I tests the intervention with a small group to evaluate the treatment's safety and potential side effects
- Phase II tests the intervention on a larger group to evaluate treatment effectiveness and further evaluate safety
- Phase III tests the intervention on a large group to confirm effectiveness, monitor side effects, and collect information that will allow the treatment to be used safely
- Phase IV is done after the treatment has been marketed to gather information about effects in various populations and the safety of long-term use

In 2013, the National Institutes of Health (NIH) granted \$2.9B for research and the pharmaceutical industry spent \$28.5B on Phase I-IV clinical trials.^{3,4} The Food and Drug Administration Modernization Act (FDAMA) of 1997 mandated registration of all Phase II-IV clinical trials, and the NIH subsequently created the ClinicalTrials.gov registry, which has been more generally adopted since 2005.⁵

Currently there are over 120,000 interventional trials in the registry, which is structured to collect detailed information regarding study design, institutional characteristics, and study results. Although investigators are beginning to analyze the database, its size and scale have led to an incomplete understanding of the current clinical trials landscape.⁶⁻⁸ Investigators have explored overall trends,⁶ and some have focused on specific specialties or diseases,⁹⁻¹¹ but these have all been superficial descriptive analyses that lack any visualization that has led to new insights.

Objectives

Our objective was to develop an interactive visualization tool which would allow users to explore the ClinicalTrials.gov registry and better understand the clinical trials landscape. The visualization was developed to allow comparisons in the number of trials conducted and patients enrolled between different disease categories, and to compare the sponsor distribution, study phase, and completion status with each study grouping. The target user is someone with pre-existing interest in clinical trials and entry-level domain knowledge, such as clinical investigators, policy analysts, funding agencies, and journalists.

Data

We obtained data for this project from ClinicalTrials.gov. This data set contains detailed information about every observational and interventional study registered with NIH (160,000 studies worldwide, of which approximately 120,000 are interventional trials). The Clinical Trials Transformation Initiative (CTTI) prepares a data extract called the Aggregate Analysis of ClinicalTrials.gov (AACT), which we obtained as a flat file and used to populate a relational database mirroring the original structure.

Tools

We imported CTTI data into a MySQL instance on the I School server and performed some initial exploratory data analysis and data cleaning in the database. We then linked the database to Tableau for visual exploratory data analysis and rapid prototyping. We used Python to extract and transform selected data into JSON objects for use in the interactive visualization, which was developed using D3.js. A GitHub repository was used for collaborative development (<https://github.com/jasonost/clinicaltrials>).

Results

The completed visualization can be accessed at <http://groups.ischool.berkeley.edu/clinicaltrials>

Upon loading the visualization, the user is presented with a help overlay that explains the main interaction mechanisms and components of the visualization (figure 1). This can be accessed at any time using the help icon in the upper right hand corner. The initial bubble diagram is organized based on top-level MeSH (Medical Subject Heading) terms and can be navigated using the defined MeSH hierarchy (figures 2 and 3). A navigation section on the top left describes the current view (total number of trials and patients enrolled), along with breadcrumbs showing the current location within the hierarchy and which selections have been made (figure 4). A link takes users to the ClinicalTrials.gov registry, showing details of all trials based on the user's current selection.

The main bubble visualization is complemented by a series of charts that provide additional trial details and comparisons. Each chart compares the current selection characteristics to all interventional trials in the registry. This can be viewed by phase of study, study enrollment and completion status, primary and secondary financial sponsor, continent where the trial was performed, and temporal trends.

Usability Testing

We fielded a survey (Appendix 1) among a select group of investigators, physicians, and developers, with eight responses (Table 1, Appendix 2). Overall, users did not have extensive knowledge of the MeSH hierarchy (mean score of 1.75, where 1 is "not at all familiar" and 5 is "very familiar"). Users felt it was an easy interface to use and understand and allowed them to explore ideas and answer questions, but they felt it was difficult to discover anomalies in the data and make comparisons. In detailed constructive comments, users felt it was a complex interface that required an investment of time and focus to extract insights. In addition, they felt the histogram allowing them to select trials based on year was not intuitive.

Future Directions

We plan to continue developing this tool, as there is a definite current and continuing need to understand the clinical trials landscape, with a large potential user base. The current state of the art consists of simple database counts and tabulations which do not provide any actionable insight. Our usability testing and feedback indicate many opportunities to make the tool more intuitive to navigate and understand, and easier for users to make comparisons between their current selection and all trials. In addition, we hope to organize the bubbles in a more meaningful and predictable fashion, better integrate with the ClinicalTrials.gov registry, and implement a search function. We hope to release this project as a freely available tool for anyone to use.

Work Breakdown

This was a collaborative project, and our individual contributions are detailed in table 2, below.

Table 1: Usability survey results

Responses to these questions were on a 1 to 5 scale, where 1 indicates strong disagreement and 5 indicates strong agreement with the prompt.

Prompt	Mean Response (SD)
It is easy to discover anomalies in the data	2.75 (0.53)
The app provides information I want to know about clinical trials	3.57 (0.53)
It is easy to discover trends in the data	3.75 (0.81)
The interface is visually appealing	3.86 (0.38)
The interface is easy to use	3.88 (0.69)
The interface is easy to understand	4.29 (0.49)
The app allows me to explore ideas and answer questions	4.43 (0.53)

Table 2: Work breakdown

Task	Colin	Anobel	Jason
Obtaining and loading the data in MySQL	10%	10%	80%
Exploratory data analysis	15%	70%	15%
Designing the visualization	40%	40%	20%
Implementing the visualization in D3	30%	10%	60%
Designing and fielding the usability study	60%	40%	0%
Documentation and presentations	20%	60%	20%

Figure 1: Initial help screen when loading visualization

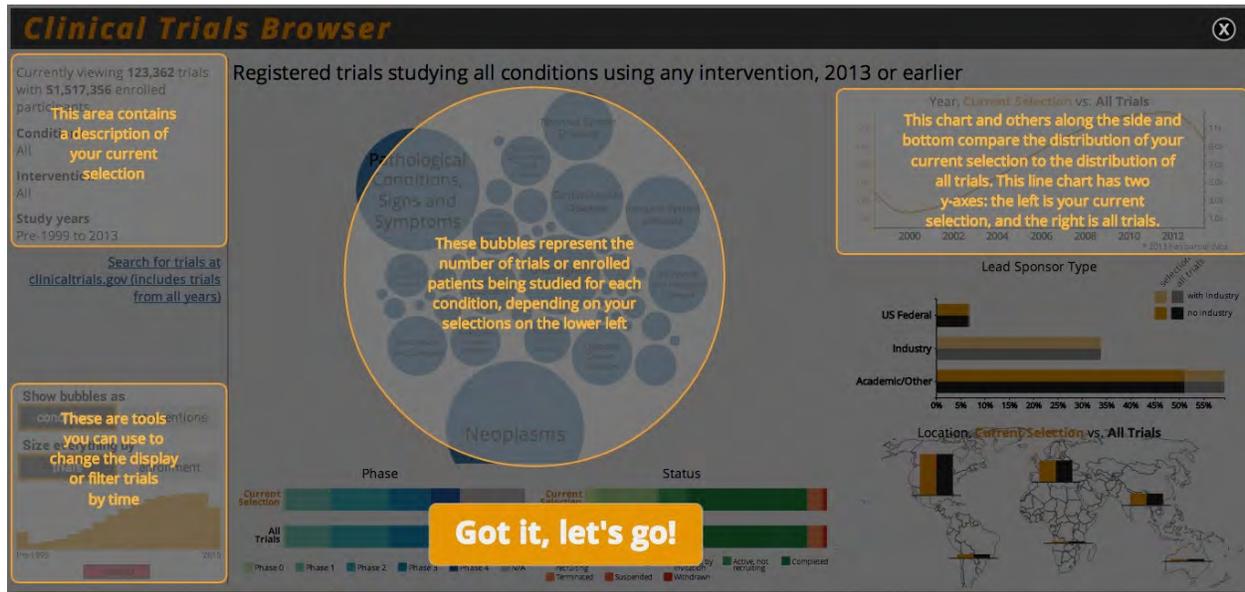


Figure 2: Clinical Trials Browser exploratory analysis tool

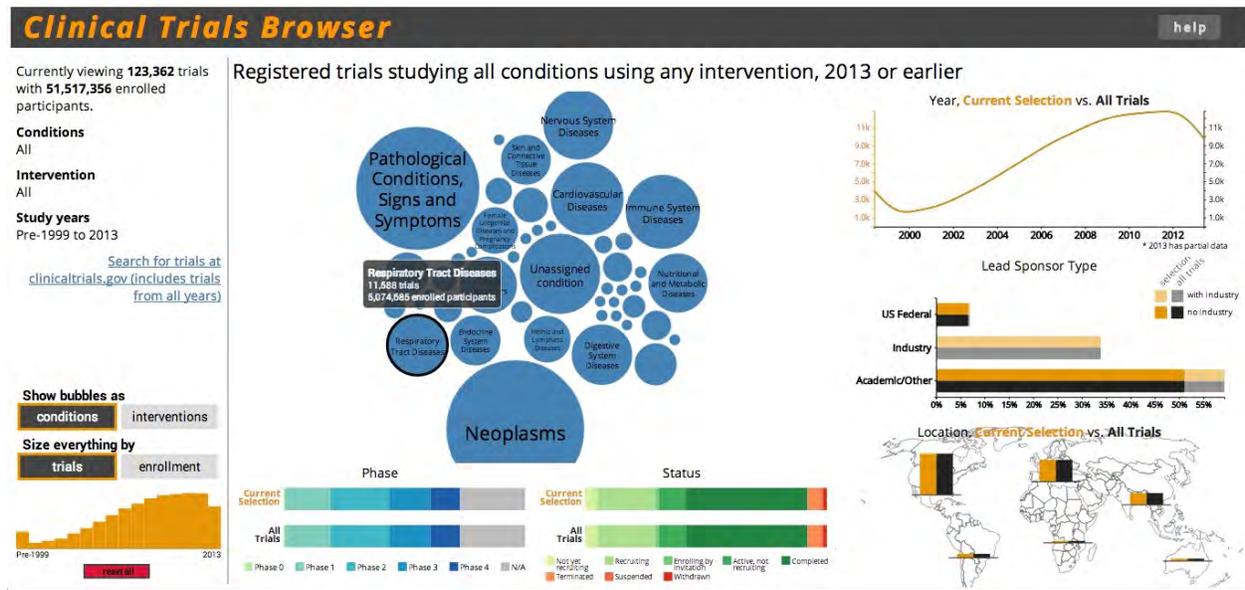


Figure 3: Navigation through clinical trial hierarchy

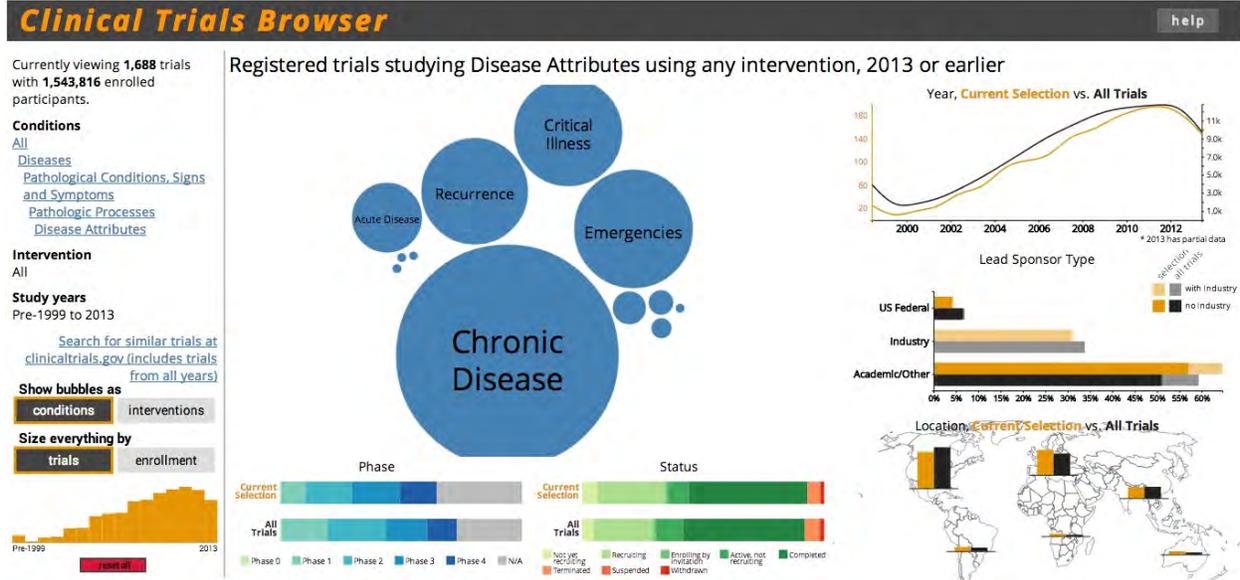


Figure 4: Navigation history with current status and breadcrumbs

Currently viewing **13,262** trials with **5,257,859** enrolled participants.

Conditions

All

[Diseases](#)

[Pathological Conditions, Signs and Symptoms](#)

[Pathologic Processes](#)

Intervention

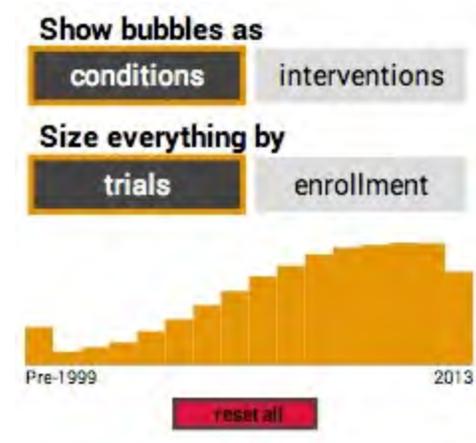
All

Study years

Pre-1999 to 2013

[Search for similar trials at clinicaltrials.gov \(includes trials from all years\)](#)

Figure 5: Options allowing user to view bubbles representing trials as the type of conditions studied or the types of interventions used, to proportionally size bubbles and other charts based on the number of trials or the number of patients enrolled, and to select specific time periods.



References

1. Learn about Clinical Studies. Available from:
<http://www.clinicaltrials.gov/ct2/about-studies/learn>
2. What are clinical trial phases? Available from:
<http://www.nlm.nih.gov/services/ctphases.html>
3. National Institutes for Health, editor. Funding Facts - NIH Research [Internet]. [cited 2014 Apr 6]. Available from: <http://report.nih.gov/fundingfacts/fundingfacts.aspx>
4. PhRMA. 2013 Profile Biopharmaceutical Research Industry. Washington, DC: Pharmaceutical Research and Manufacturers of America; 2013.
5. Roumiantseva D, Carini S, Sim I, Wagner TH. Sponsorship and design characteristics of trials registered in ClinicalTrials.gov. *Contemp Clin Trials* 2013;34(2):348–55.
6. Califf RM, Zarin DA, Kramer JM, Sherman RE, Aberle LH, Tasneem A. Characteristics of clinical trials registered in ClinicalTrials.gov, 2007-2010. *JAMA: The Journal of the American Medical Association* 2012;307(17):1838–47.
7. Quesy SN. How should we interrogate the ClinicalTrials.gov database? *J Pain* 2013;14(9):997–8.
8. Riveros C, Dechartres A, Perrodeau E, Haneef R, Boutron I, Ravaud P. Timing and completeness of trial results posted at ClinicalTrials.gov and published in journals. *PLoS Med* 2013;10(12):e1001566–discussione1001566.
9. Witsell DL, Schulz KA, Lee WT, Chiswell K. An analysis of registered clinical trials in otolaryngology from 2007 to 2010: ClinicalTrials.gov. *Otolaryngology -- Head and Neck Surgery* 2013;149(5):692–9.
10. Goswami ND, Pfeiffer CD, Horton JR, Chiswell K, Tasneem A, Tsalik EL. The State of Infectious Diseases Clinical Trials: A Systematic Review of ClinicalTrials.gov. *PLoS ONE* 2013;8(10):e77086.
11. Inrig JK, Califf RM, Tasneem A, et al. The Landscape of Clinical Trials in Nephrology: A Systematic Review of ClinicalTrials.gov. *Am J Kidney Dis* 2013;

Clinical Trials Browser Questionnaire

This is a questionnaire for giving feedback on our prototype web app. We are looking for your honest feedback on which features you like and which features are difficult to understand or use.

This app enables people to explore the clinicaltrials.gov database of trials registered with the National Institutes of Health (NIH). It consists of more than 160,000 clinical studies that have been performed over the last 15 to 20 years. All of the studies are tagged with terms from the Medical Subject Headings (MeSH terms) that we used to enable exploration. More information about MeSH terms can be found at http://en.wikipedia.org/wiki/Medical_Subject_Headings

Please click this link to explore our app while going through the questions:
<http://groups.ischool.berkeley.edu/clinicaltrials/> (NB: the app works best in Chrome or Firefox on a desktop or laptop computer. It may not work on a phone or tablet.)

This app is an early prototype which will likely have bugs and features that do not function as expected. We would really appreciate it if you could point out any inconsistencies you notice.

You can click on the blue bubbles to "drill-down" into the MeSH hierarchy of conditions studied in clinical trials.

Was this obvious? Can you tell where you are in the hierarchy? How to go up one or more levels? Please include any other thoughts or comments about how the interface allows you to explore this hierarchy.

You can click on the conditions/interventions or trials/enrollment option buttons in the lower left to change the display.

Are these useful? Are they easy or difficult to understand?

You can select a time range using the bar chart in the lower left (above the reset button).

Was this obvious? Does this work as you expected?

When using the Clinical Trials Browser app, how much do you agree with each of the following statements?

	Strongly disagree	Somewhat disagree	Neutral	Somewhat agree	Strongly agree
The app allows me to explore ideas and answer questions	<input type="radio"/>				
The app provides information I want to know about clinical trials	<input type="radio"/>				
It is easy to discover anomalies in the data	<input type="radio"/>				
It is easy to discover trends in the data	<input type="radio"/>				
The interface is visually appealing	<input type="radio"/>				
The interface is easy to use	<input type="radio"/>				
The interface is easy to understand	<input type="radio"/>				

Before using this tool, how familiar were you with the MeSH hierarchy?

1 2 3 4 5

Not at all familiar Very familiar

What types of questions could you answer with this tool?

What parts of the site did you find especially useful?

Which parts of the site did you find difficult to use or understand?

This question is especially important to us.

Please write any additional comments you have about the app.

Submit

Never submit passwords through Google Forms.

Powered by
 Google Drive

This content is neither created nor endorsed by Google.

[Report Abuse](#) - [Terms of Service](#) - [Additional Terms](#)

Appendix 2. Open-ended Survey Responses

Overall constructive feedback

- Overall it needs work allowing users to drill down along the specific dimensions they care about
 - I also don't know if you need to have that many charts on one page.
- If I were to do this:
 - I would have a super clean interface for easily selecting dimensions you want to explore. Clearly showing you the hierarchy and how many different trials are available under each.
 - I would then on the right hand side let the user pick the type of info/chart they are interested in. You can look at Google Analytics for a bit of inspiration here
 - I would add search
- A good amount of the information on the page required me to mouse over in order to get useful information (e.g., the smaller, unlabeled bubbles or the "Trials by Study Phase" and "Trials by Status" areas). While it is good that the information was available, I found that there was slightly too much reliance on mouse overs.
- Overall, I felt that there was a lot of good information on the site, but I experienced a bit of information overload. I would recommend considering changing the interface such that only one data display is shown at a time. The hierarchy and filters can still be shown in all views, but a tab menu would enable navigation between the different data displays. This might make it easier for people to focus on the information at hand while still giving them easy access to all the data options."

Usability of bubble navigation

- I discovered clicking on the bubbles very quickly. Several levels down I started to feel lost but then I noticed the hierarchy listed on the left. Still, I was left with the sense that I should be rewarded for clicking on all those bubbles--i.e., where's the pot of gold at the end of the rainbow?
- The bubbles sit in the main frame so it's not obvious that it's clickable, given that the others aren't - perhaps having them in a panel that makes it obvious that it's also part of the controls to browse the dataset would help.
- It took a few minutes of clicking to realize where to look to find where I was in the hierarchy - I was looking for breadcrumbs, a back or 'up' icon in the vicinity of the bubbles, but later realized that the hierarchy is reflected on the right as you drill down. A glow or animation of some sort in the section on the left to draw attention to the change after the bubbles settle would point that out well.
- Drill down a ways having bubble with category are already in but smaller - didn't understand. Eg. all study years, bubbles by conditions sized by trials: all>diseases>neoplasms>neoplasms by histologic type>neoplasms, glandular and epithelial had a bubble for 'neoplasms, glandular and epithelial' - is this meant to be 'other'?
- Clicking on the bubbles was reasonably obvious
- The hierarchy is confusing however. Typically a navigation for a hierarchy is done in the form of breadcrumbs UI element (<http://ui-patterns.com/patterns/Breadcrumbs>) but here its' done in the lefthand nav in a way that was not obvious to me at first"
- Yes, I can click on the bubbles. It wasn't obvious that this provides a drill down.

- Maybe informing in hover tool tips that you can select the category would help.
- Takes a few minutes before you get the hang of it. the hardest part was getting back to the top. didn't find the "reset" button in the bottom left corner initially.
- Yes, it was obvious to anyone used to using these types of visualizations... however anyone less tech savvy might not understand that the tabbed hierarchy that shows up on the left side of the page allows you to move back up the hierarchy...
- It was intuitive that one could click on the blue bubbles to "drill-down" in the hierarchy.
- During my first session on the site, I did not notice the feature that showed me where I was in the hierarchy and enabled me to go up one or more levels. After viewing this question, however, I went back to the site, and I was easily able to find the hierarchy listing in the top left corner.
- Good layout, easy to tell relative levels/numbers by bubble size

Usability of brushing interaction

- I discovered this feature without being prompted, it is very interesting, however, it is difficult to assess the impact on the bubble view before and after terms. Perhaps a better animation that shows how things change as you slide the time slider, rather than resetting the bubbles completely with each update?
- Not obvious, no - neither was being able to drag to select multiple bars, just gave it a try whilst exploring the ui for this questionnaire
- This is a very poor UI element and not at all obvious. I think you should copy a simple date range UI element. Just use travel sites such as Kayak and the likes for inspiration
- Yes, I love the d3 brush graph. Its very smooth, but it is definitely not very intuitive that it is capable of selection vs. a static bar chart.
- One thing that would help, make the brush selection contain the entire time horizon on the default or "reset" view. Currently there is no selection on the chart initially. By adding an extra year at the beginning or the end of the time series and having the brush selection initialize to a subset, it may make this even more clear.
- not very obvious. like i said earlier it takes a few mins before you get the hang of things. but it does work as expected
- This was less clear. if you could have a popup when you scroll over it that gives instructions that could help. Also, an additional feature if you have time would be to compare years. So what does 2004 compare to 2013.
- This was not obvious. I discovered this feature when I moused over that area of the display and the cursor changed. It might be more intuitive to have little, drag-able carrots along the year axis.

Components that were difficult to understand

- Putting a world map on there was overkill (i.e. busy and distracting) when it didn't actually break things down by country, just by continent. If you're just breaking down by continent a bar chart or table would suffice.
- How to narrow down my selections based on available dimensions (conditions, date, etc)
- Some of the UI was a bit confusing about what could be clicked and what couldn't.

- Having an introduction to the interface could be really helpful. You could place semi-transparent overlays over each component to highlight the functionality for selecting time, category, or region.
- The bottom area showing the "Trials by Study Phase" and "Trials by Status" were too small to read easily, and it made me want to avoid taking the time to figure out what they were. The ability to mouse over different segments for more detailed information was good, though.

Components that were especially useful

- Drilling down into diseases and getting an exact link to the clinicaltrials.gov site
- I'm not a medical professional...just providing friendly feedback.
- This site can be very helpful even for the lay user. As a person who does not actively participate in research in any medical field, i still found the site to be a useful way to explore what type of research is currently being performed.
- The bar graph showing the percentage of trials with a lead sponsor type gave me good information about what kinds of organizations are contributing to research in a particular clinical trial topic. This was one of the more intuitive data displays.