

# ClinicalTrials.gov as a data source for point-of-care trial recruitment



PB Pfiffner, J Oh, TA Miller, KD Mandl

pascal.pfiffner@childrens.harvard.edu — <http://p2.github.io/clinical-trials-eval>

141

## many trials, low enrollment

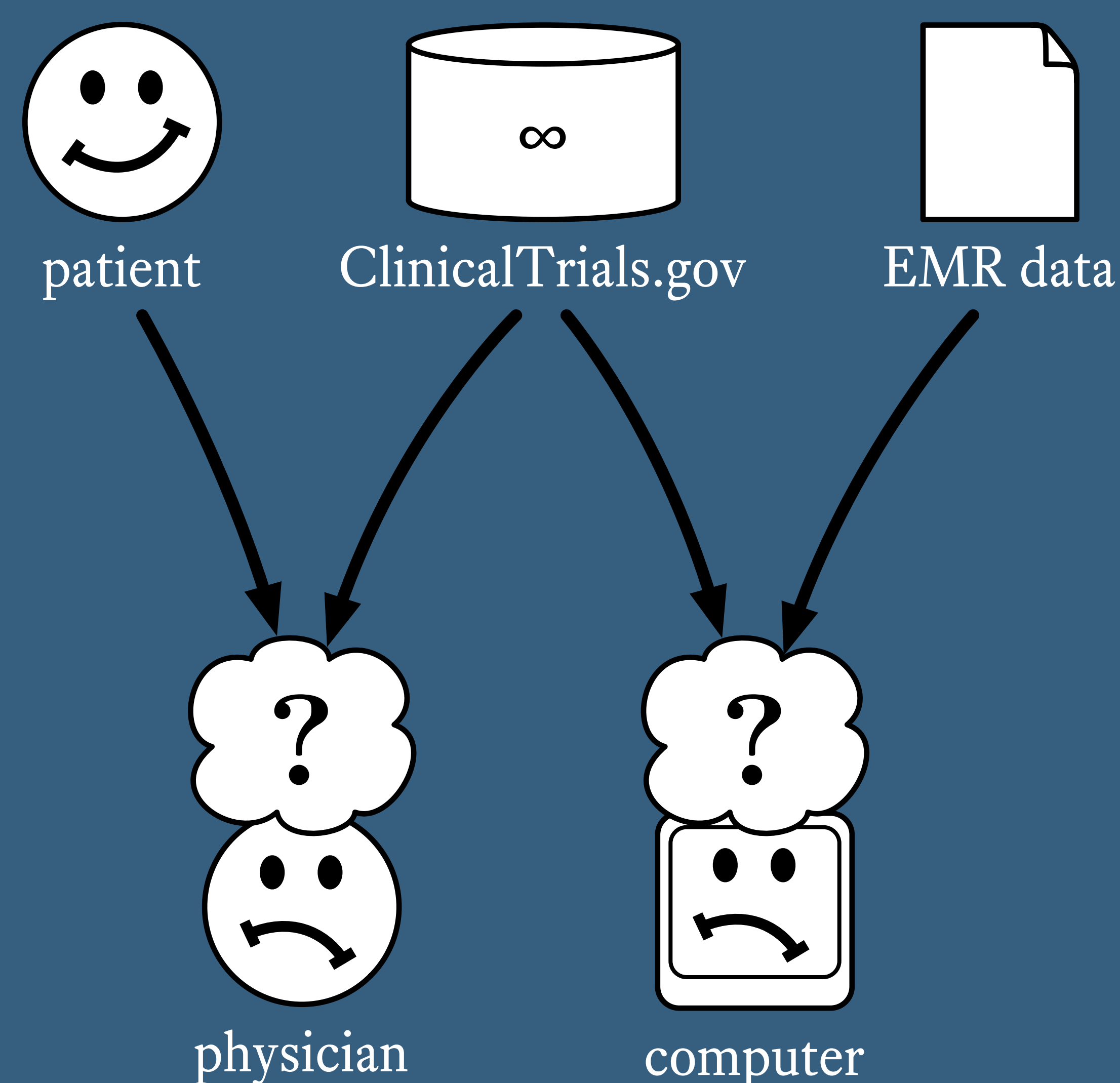
With over 50,000 active trials registered in ClinicalTrials.gov (CTG), finding the right trial is a laborious task and enrollment targets are often missed. Armed with EMR data, the physician could play a key role in identifying and recruiting patients at the point of care. However, eligibility criteria in CTG are provided as free text and thus not “readable” by a computer.

Can we leverage our investment into EMRs and build apps that use patient data, for example via the SMART platforms project, to enable automated trial recruitment at the point of care?

## data quality?

To determine to which extent computational tools might assist in eligibility screening we evaluated CTG trial data by asking:

- is the trial’s recruitment status up to date?
- are there significant barriers to computational eligibility criteria extraction?
- is the contact information correct?



## methods

Download trials with an “open” recruitment status for 4 search terms.

437	98 Gleevec
	113 cataract
	124 neuroblastoma
	102 rheumatoid arthritis

Manually and computationally evaluate:  
calendar dates  
eligibility criteria  
contact information

Follow-up by phone and email.

146	30 Gleevec
	39 cataract
	40 neuroblastoma
	37 rheumatoid arthritis

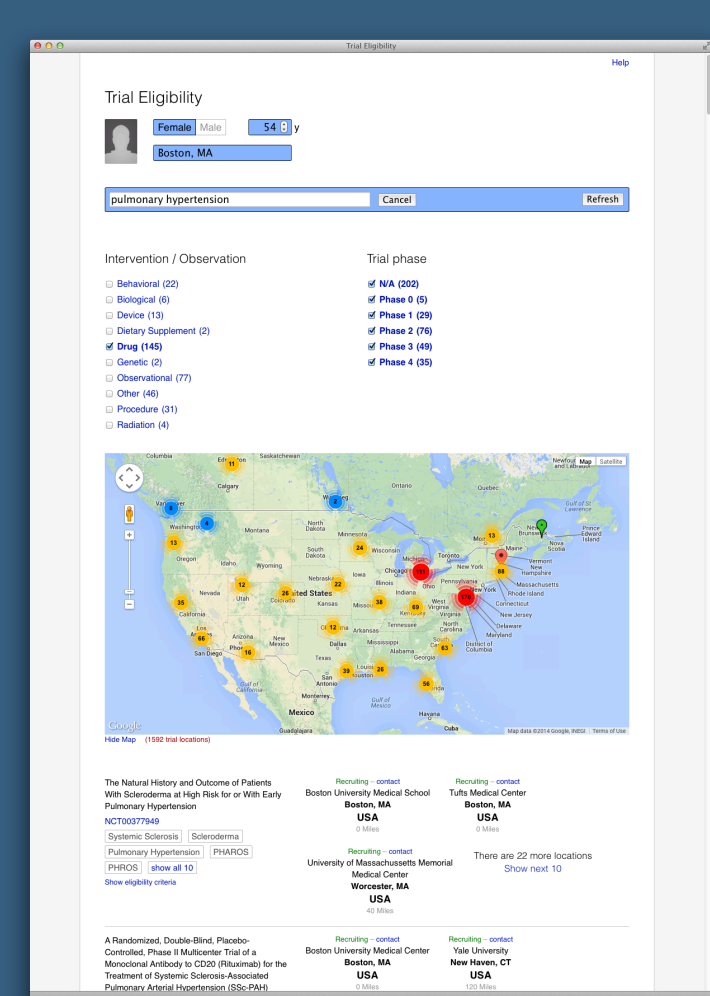
## results

14 - 46%  
Outdated recruitment status

97 - 99%  
Standard eligibility criteria text format

81 - 94%  
Significant barriers to automated eligibility criteria extraction

31%  
Could not be contacted, neither by phone nor by email



top 5

Can a computer provide a top 5 list and why don't we yet have it?

## low quality

Registry data are partially out of date or inaccurate. Eligibility criteria are largely unstructured and therefore not readily amenable to automated trial eligibility determination.

## NLP

Using Natural Language Processing (NLP) for eligibility criteria extraction holds promise but considerable obstacles must be overcome.

## key criteria

Entering a handful of key eligibility criteria into CTG in structured form might enable very powerful initial, automated trial filtering.