Comparative Strengths of Public and Commercial Clinical Trials Databases: A Case Study

Authors: Diane Webb, Matt Eberle and John Willmore – BizInt Solutions, Inc.   **Objective**: In this case study, commercial and public clinical trials databases were searched for chosen diseases in order to evaluate differences in trial coverage and content.

Introduction

If you are looking for information on ongoing and completed clinical trials, which databases should you use? What value do you get from working with more than one source of published clinical trial data? Given the wealth of clinical trials available in ClinicalTrials.gov, is there additional insight to be gained by adding sources like EU Clinical Trials (EudraCT), WHO International Clinical Trials Registry Platform (ICTRP), Citeline TrialTrove, Thomson Cortellis Trials Intelligence and Adis Clinical Trials Insight?

In this case study, we will search several commercial and public clinical trials databases for selected diseases and evaluate differences in trial coverage and content. The case study will illustrate how elements of unique content in each database can be used to support competitive analysis and trial planning.

Method

We searched six trial databases (three commercial and three public) using database-specific terminology where available, for two smaller oncology indications (mesothelioma and macroglobulinemia) and pertussis. The following clinical trial databases were searched:
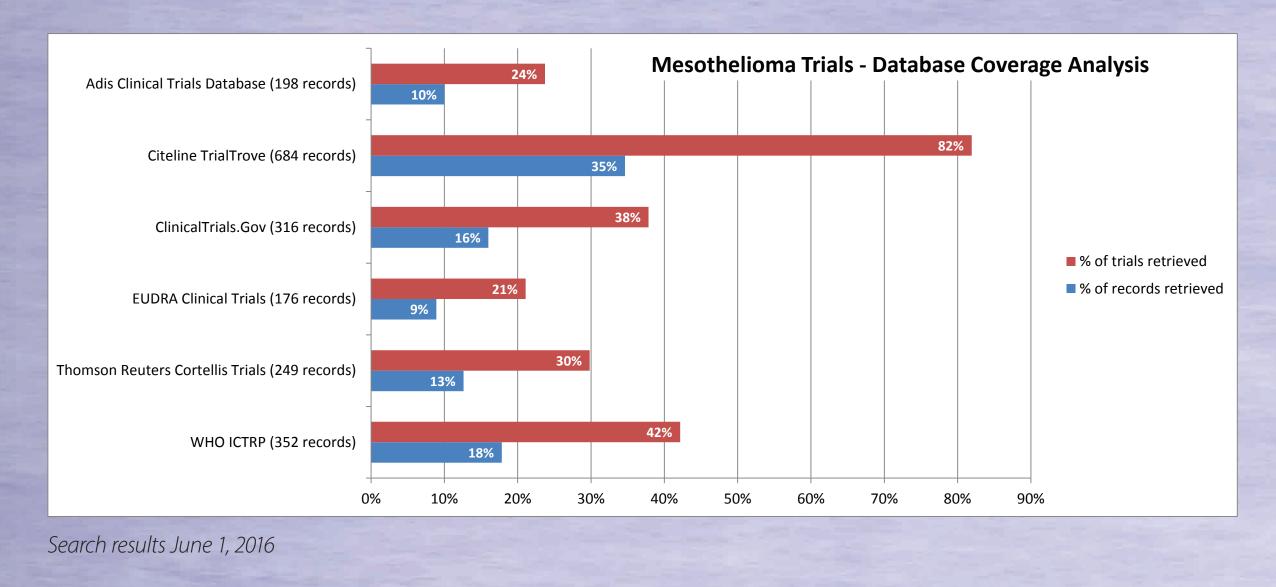
- NIH ClinicalTrials.gov
- European Union EudraCT
- World Health Organization ICTRP
- Citeline TrialTrove
- Adis Clinical Trials Insight
- Cortellis Trials Intelligence.

The records retrieved were combined into a single report for each indication and the "Identify Common Trial ID" tool was used to match related trials across databases. Statistics were calculated for the percentages of total *records* and total *trials* retrieved from each database.
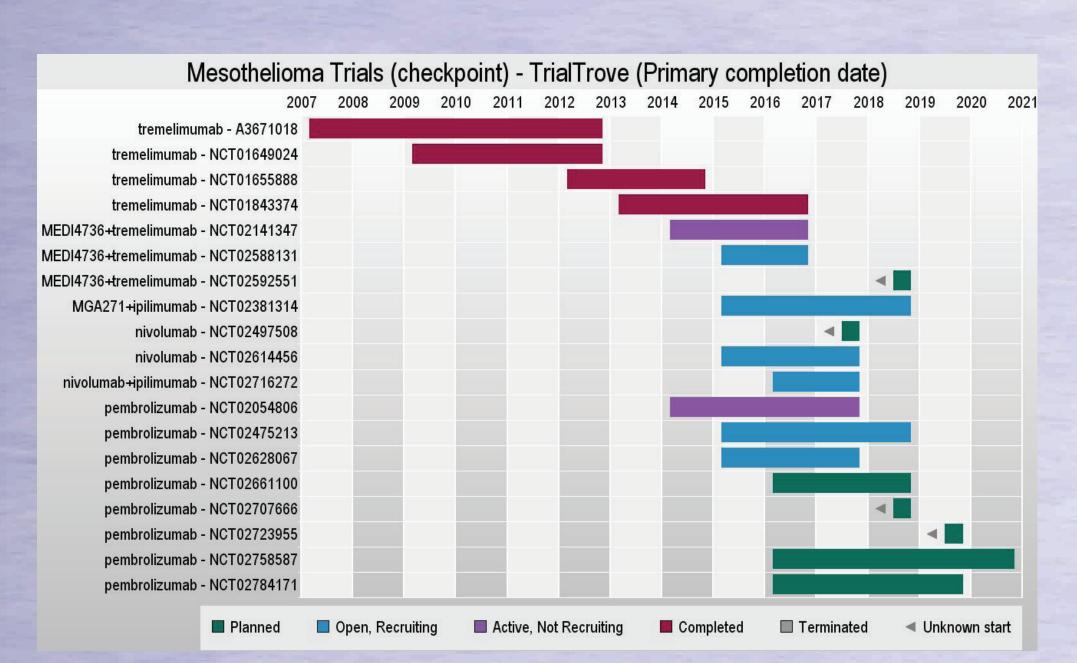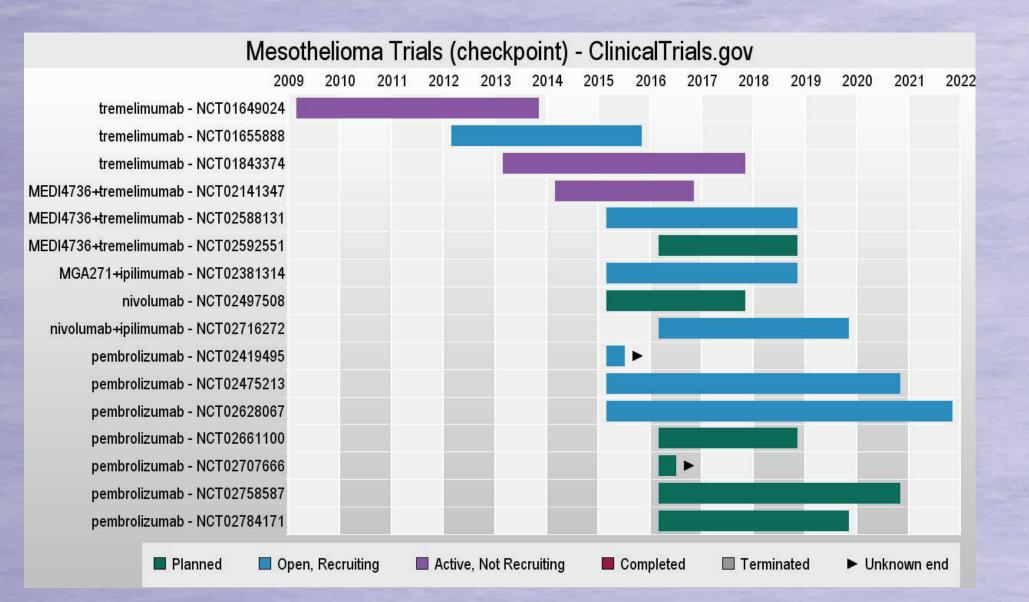
We then created trial timelines for a selection of mesothelioma checkpoint trials (from both single databases and from the combination of all six databases) to evaluate coverage and content variation.

Results

Mesothelioma: 1975 total records were retrieved, representing 835 trials.

Search results June 1, 2016.

Macroglobulimenia: 1806 total records were retrieved, representing 958 trials.

Search results June 1, 2016.

Pertussis: 1586 total records were retrieved, representing 551 trials.

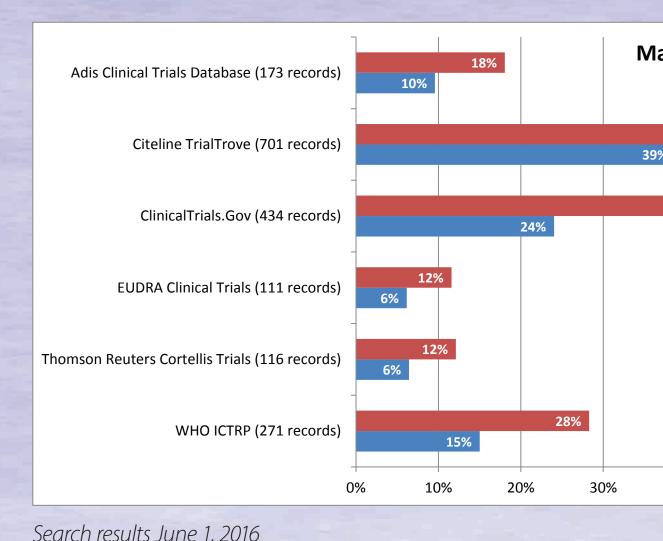Search results October 2, 2015.

Comparing TrialTrove and ClinicalTrials.gov: Mesothelioma Checkpoint Inhibitors Trials
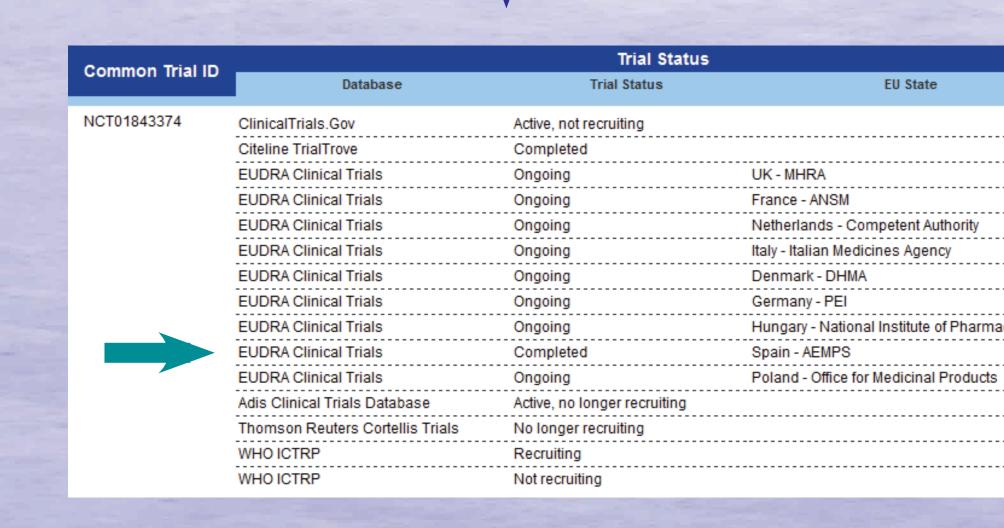

Four trials (A3671018, NCT02614456 , NCT02054806, NCT02723955), were retrieved from TrialTrove but not from ClinicalTrials.gov. A3671018 was retrieved only from TrialTrove (out of all 6 database searches).

One trial (NCT02419495) was retrieved from ClinicalTrials.gov that was not retrieved from TrialTrove. Two trials not retrieved from ClinicalTrials.gov (NCT02614456 and NCT02723955) were retrieved from Cortellis, providing completion dates missing in TrialTrove. These trials (and the completion dates) could also be retrieved from ClinicalTrials.gov by searching the NCT numbers.

The latest starting trial for tremelimumab, NCT01843374 is shown as active (no longer recruiting.) CT.gov, Adis CTI, and Cortellis all show this status and have the most recently updated records; TrialTrove shows the trial as complete. Looking at the EudraCT data, we see that only the trial record for Spain shows the trial as completed.
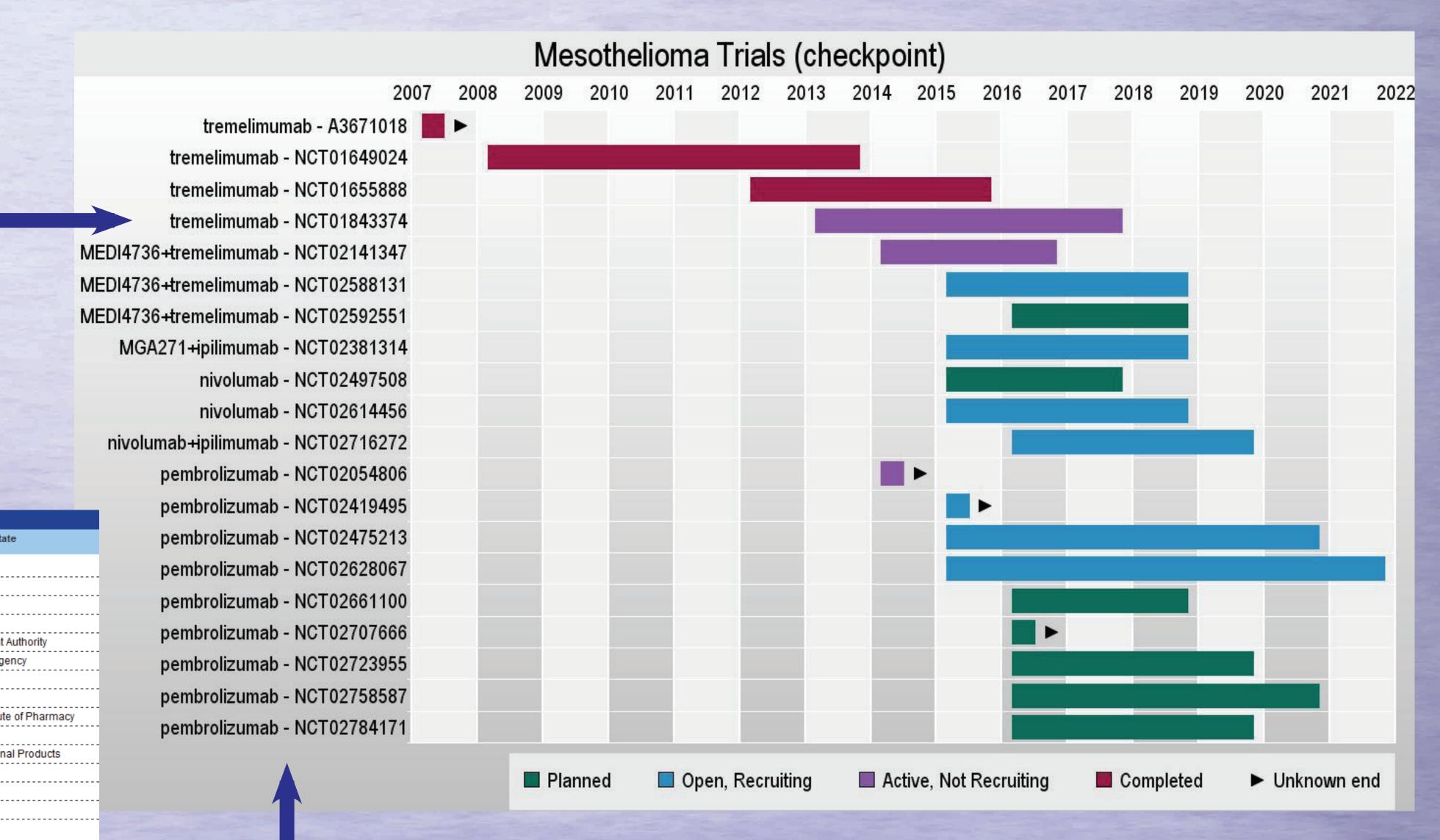

Mesothelioma Checkpoint Inhibitors Trials: Leveraging data from all trial databases

This trial timeline leverages data from all six trial databases by selectively integrating each trial and each timeline element (e.g. trial status, start date, end date.) For example, for each trial we display the earliest start date from any source and completion dates from CT.gov. The trial status was selected from the most recently updated trial record.

For pembrolizumab, CT.gov picks up an early trial (NCT02419495) that TrialTrove does not. Other sources gives us completion dates for the next two pembrolizumab trials that extend beyond the primary endpoints given by TrialTrove. TrialTrove picks up NCT02723955 which wasn't retrieved in the CT.gov search. But, TrialTrove had neither start nor completion dates for these trials.

Conclusions

For the two oncology searches, the Citeline TrialTrove search retrieved 70-80% of the trials retrieved from all six databases. For the pertussis search, the public databases, especially ClinicalTrials. gov and WHO ICTRP, retrieved 70-80% trials. Adis CTI also retrieved a large number of trials (61%.) Pertussis is not a disease area covered by Citeline TrialTrove, which retrieved only 13% of trials.

In order to create the mesothelioma checkpoint trial timeline, we needed to supplement Citeline TrialTrove coverage and content with data from other trial databases, especially for trial completion dates.

Our experience with clinical trial (and drug pipeline) data over two decades confirms that "duplicate" data is a misnomer. Databases covering the same topic display differing—sometimes surprisingly so—strengths and weaknesses by region covered, therapeutic area, vocabulary standardization, update frequency and more.

Each of the commercial and public clinical trials databases have certain strengths in coverage and content. Some databases provide excellent information for a specific country or region, while others provide global coverage. Commercial databases tend to focus on key therapeutic areas. Public databases can provide better coverage for rare diseases or public health concerns. Finally, database indexing policies and update frequency can result in content differences between data for the same trials.

Disclosure