Poster presented at DIA 2018

For more information:
contact support@bizint.com

Software for Business Intelligence
BizInt Smart Charts
www.bizint.com

Visualizing Clinical Trial Endpoints

Author: Matt Eberle — BizInt Solutions, Inc.

Objective

Can endpoint terms be extracted and normalized from clinical trial registries to build visualizations to facilitate trial planning and competitive intelligence?

Method

We searched both US and EU registries for two diseases, Merkel Cell carcinoma (MCC) and progressive multiple sclerosis (MS) and combined the data into a single table for each disease. For comparison, we ran a search for both diseases in Cortellis Clinical Trials Intelligence (CTI) from Clarivate Analytics.

Table 1. Search results from US & EU registries and Cortellis CTI for selected indications.

	US & EU Registries		Cortellis CTI
	# of records	# of trials	# of trials
Merkel Cell Carcinoma	77	51	60
Progressive Multiple Sclerosis	534	229	358

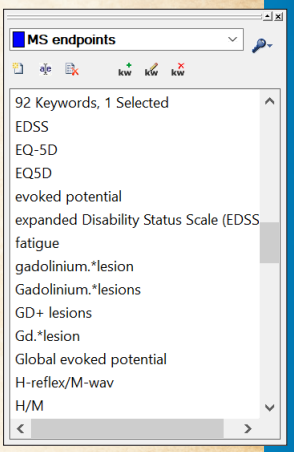
By reviewing the retrieved records and published literature, we identified a set of endpoint terms and used text-mining tools to extract these terms from registry trial records. We then built a thesaurus to normalize endpoint term variations. In Cortellis CTI endpoints are assigned controlled terms in addition to free text, simplifying extraction and normalization.

Using the normalized terms from each source, we then created visualizations (see examples on pages 2-3) to facilitate analysis.

See pages 2-3 for trial endpoint visualizations.

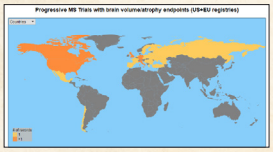
Results – Term Extraction

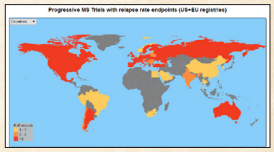
Endpoints or outcomes in the US and EU trial registries are not entered using a controlled terminology. Even with established endpoints we found variations in wording across trial records and between different registry records for the same trial. For trials phase 1-3 we identified at least one primary endpoint per trial for 88% of MCC trials and 89% for progressive MS trials.

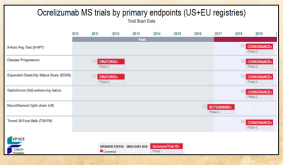


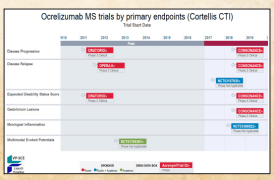
We found similar endpoint concepts in both extracted terms from registry records and indexed terms from CTI. The ontology in CTI revealed additional categories of interest, such as Patient Report Outcomes. And, we found unique records in each source. So, a combination of controlled terminology and term extraction will yield the most complete picture.

The visualizations included here are selected examples illustrating the potential of extracted and indexed endpoint terms.









Results – CTI Indexing

Recently introduced indexing enhancements in Cortellis Clinical Trials Intelligence (CTI) offered several benefits in this study.

Improved indexing of indications in CTI resulted in increased recall in both of the indications we investigated, as seen in Table 1 on page 1. When comparing trials retrieved from CTI to the results from the registries, the columns labeled "# of trials" should be compared. The number of records indicated for the registries is larger, indicating overlap between the US and EU registries, as well as the presence of records for more than one EU state in many trials.

The use of a controlled vocabulary in CTI endpoint indexing meant we did not need to extract keywords from free text (as required from registry data), simplifying the process of term extraction.

Endpoint index terms in CTI can have two levels. For the visualizations here we selected only the first level terms. This allows for broader trends to be identified that aren't always visible with the terms extracted from the registry records.

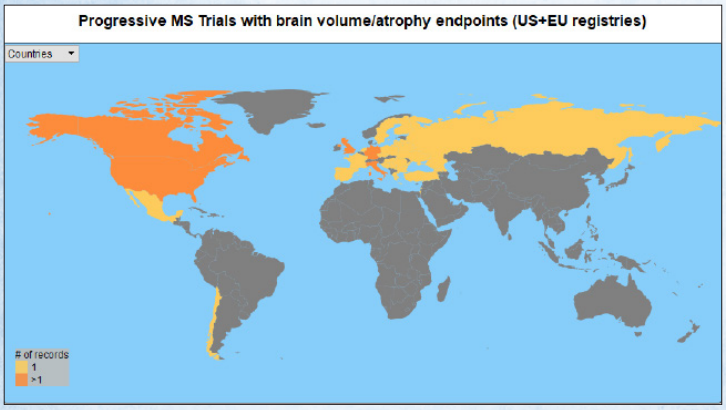
Cortellis Trials Intelligence (CTI): Progressive MS

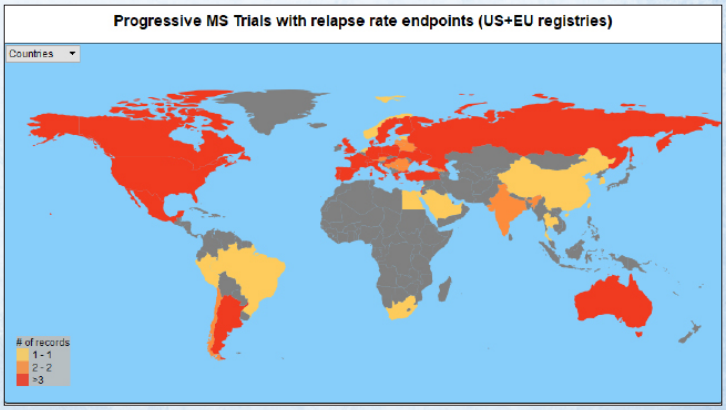
	Primary Endpoints	Secondary Endpoints
1	Assessment of Disease Relapse/Recurrence Assessment of Disease Relapse/Recurrence - Annualized relapse rates	Assessment of Other MRI Variables
2	Assessment of Other MRI Variables Assessment of Other MRI Variables - Brain atrophy	Assessment of Disease Progression
3	Assessment of Multiple Sclerosis Functional Composite (MSFC) Score Assessment of Multiple Sclerosis Functional Composite (MSFC) Score - Paced Auditory Serial Addition Test (PASAT)	Assessment of Neuropsychological Function Assessment of Neuropsychological Function - Stroop Color and Word Test Assessment of Neuropsychological Function - Assessment by Symbol Digit Modalities Test Score
4	Assessment of Disease Progression Assessment of Expanded Disability Status Score (EDSS)	Assessment of Disease Relapse/Recurrence Assessment of Disease Relapse/Recurrence - Annualized relapse rates Assessment of Other MRI Variables Assessment of Safety and Tolerability Assessment of Safety and Tolerability - Assessment of serious adverse events Assessment of use of Rescue Medication
5	Assessment of Safety and Tolerability Assessment of Therapy Related Outcomes Assessment of Therapy Related Outcomes - Assessment by imaging techniques Assessment of Therapy Related Outcomes - Assessment of efficacy/effectiveness of intervention Assessment of Safety and Tolerability - Assessment of adverse events/serious adverse	Assessment of Therapy Related Outcomes Assessment of Therapy Related Outcomes - Efficacy Assessment of Disease Progression Assessment of Therapy Related Outcomes - Assessment of efficacy/effectiveness of intervention

Examples of Primary Endpoint and Secondary Endpoint indexing from Cortellis Clinical Trials Intelligence (CTI).

Results – Visualizations

Maps – Where Are Extracted Endpoints Being Tested?

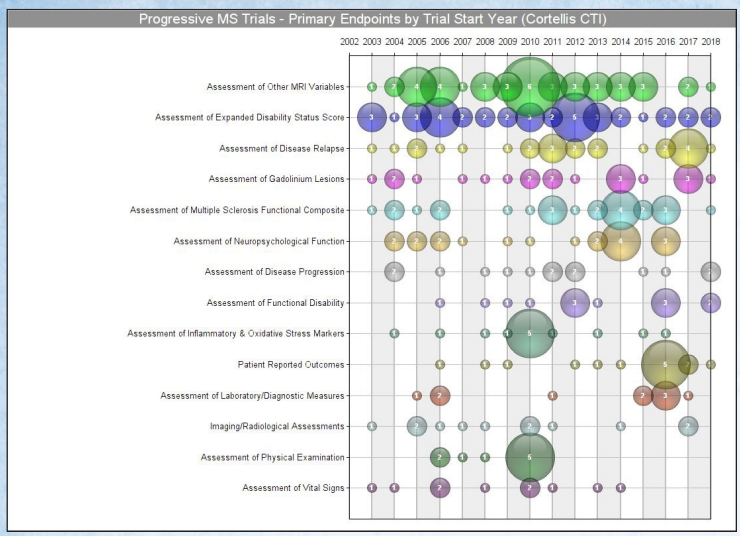


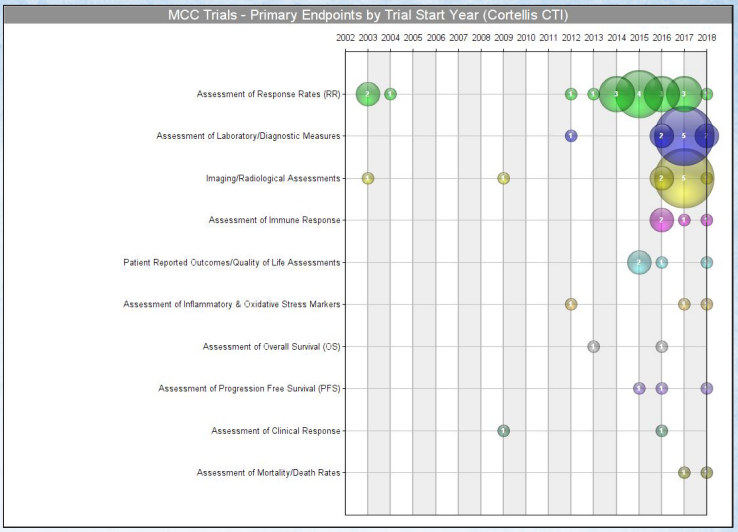


These two world maps cover nearly the same number of trials—12 trials with brain volume/atrophy as an endpoint (left map) and 11 trials with relapse rate as an endpoint (right map). But there is a big difference in geographic coverage with relapse related endpoints showing trials in Asia, Australia, Africa, and additional countries in South America.

Results – Visualizations

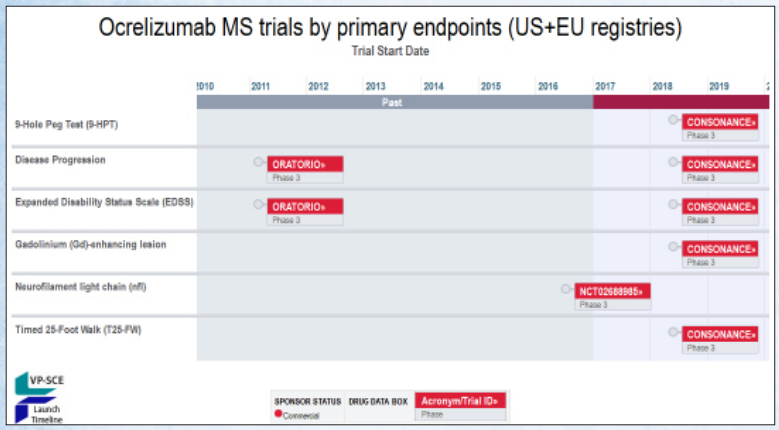
Bubble Charts – How Do Primary Endpoints Evolve Over Time?

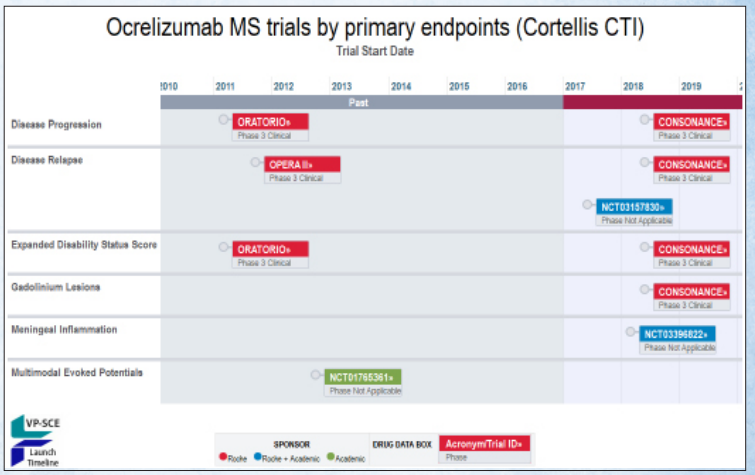




These two bubble charts show trends in primary endpoint terms extracted from controlled terminology from CTI. Similar endpoints were identified in the registry search but the categorization is different as was the number of trials, especially for the progressive MS search. Enhanced information in CTI allows for additional trends to be identified such as the increase in patient reported outcomes and the disappearance of physical examination endpoints in the progressive MS example. With the MCC search we can see an emphasis on response rates with only occasional trials using survival as primary endpoint.

Trial Start Timelines – What Endpoints Are Being Evaluated In Trials For A Selected Drug?





Using a timeline with each trial shown in multiple categories we can see multiple endpoints per trial and trends in endpoints at the trial level. Ocrelizumab was approved in the US in 2017 on the basis of ORATORIO and OPERA II. The CTI search identified a few trials not seen in the registry search and the registry search identified one additional trial (NCT02688985) not in the CTI search.

Looking at the timelines we can see the endpoints provide a strong indication of the reason for each trial. The two registrational trials focus on different primary endpoints and the post-approval CONSONANCE trial adds an additional primary endpoint.

Visualizing Clinical Trial Endpoints

Conclusions

Trial endpoints are key to both trial planning and competitive intelligence. But, the lack of controlled vocabulary for endpoints in trial registries presents a challenge for useful analysis. A manual approach requires reviewing each record individually. With software we automatically extracted identified endpoint terms.

We extracted at least one primary endpoint term from nearly 90% of phase 1-3 trial records with endpoints listed. This could be increased by adding additional terms. However, endpoints not captured were generally trial specific and would not impact a broad visualization such as the bubble chart or world map. For visualizations like the timeline that plot each trial, manual review is recommended to supplement this approach.

We successfully used software tools to identify endpoints from unstructured text and created visualizations to support analysis for clinical trial planning, regulatory strategy, and competitive intelligence. The results for MCC suggest that the same methods, keywords list, and thesaurus will be applicable to other oncology indications. The small set of trials found in the registries make it difficult to identify broad trends. But endpoints still proved a useful factor to differentiate trials for a single drug. The search for progressive multiple sclerosis yielded a larger dataset and so is better represented in the visualizations included here.

While endpoints may be different for other indications, the approach shown here is easily applicable to other indications. More importantly, the visualizations are adaptable to your key questions to better inform your strategy.

Software Tools

BizInt Smart Charts

Drug Development Suite

BizInt Smart Charts Drug Development Suite was used to import search results from ClinicalTrials.gov, EU Clinical Trials and Cortellis CTI and to build a tabular report with fields of interest from each source. These reports were then combined and the "Identify Common Trial ID" tool was used to match related trials across databases. Key data was exported to VantagePoint - Smart Charts Edition for further analysis.

vantagepoint
Smart Charts Edition

VantagePoint – Smart Charts Edition (VP-SCE) is a data analysis and visualization tool which works with data exported from BizInt Smart Charts. We used VP-SCE to identify endpoint terms, extract endpoint terms from records, and normalize terminology using a custom-built thesaurus. VP-SCE was then used to create the visualizations, including the trial timelines (which were purpose-built for pharmaceutical analysis.)

For more information and a free trial:

www.bizint.com/dia2018