



TLF Treasure Hunt: Making Legacy Data Instantly Discoverable with RAD Finder

Cielo Marie Castañeda
Pfizer

Agenda

01 **Introduction**
Background

02 **Overview**
RAD Finder App

03 **Features**
User Workflow

04 **Benefits and Impact**
Limitations and Considerations

05 **Conclusion**
References



Disclaimer

We confirm that, the opinions and thoughts discussed in this presentation are subject to our own independent views and are not subject to the opinions of the organization that we represent.

Introduction

Unprecedented surge in clinical data generation

A light blue downward-pointing arrow indicating a logical flow from the first point to the second.

Millions of data points per trial

A light blue downward-pointing arrow indicating a logical flow from the second point to the third.

Manual retrieval is time-consuming and inefficient

A light blue downward-pointing arrow indicating a logical flow from the third point to the fourth.

Need for metadata-driven tools

Background

(how did we come up creating the app?)

A diagram consisting of four dark blue rounded rectangular boxes arranged in a 2x2 grid. The boxes are interconnected by a light gray, semi-transparent background shape that resembles a four-pointed star or a cross with rounded corners. Each box contains white text describing a challenge or context in the industry.

**Thousands
of TLFs per
study**

**Critical for
regulatory
submissions
and safety
summaries**

**Manual
searches
lead to
duplication
and delays**

**Lack of
centralized,
metadata-
driven
search
capabilities**

Primary repository for finalized TLFs

Robust storage but lacks advanced search

Manual searches across systems are inefficient

Need for complementary solution

RAD Finder App



Shiny-based application



Centralized access to TLFs



Metadata-driven search



Direct links to PDFs and files

Features

01

**Metadata
Filters**

02

**Keyword
and
exclusion
search**

03

**Run date
filtering**

04

**Footnotes
search**

05

**Direct
access
links to
outputs**

FEATURES: Metadata Filters

- Users can filter by:
 - Compound Name
 - Submission
 - Protocol
- Selections can be made via dropdown menus or manual input, with an option to select “**All**” for broader searches.

Total Compounds: 25

Search by Compound Name

All

Search by Submission

All

Search by Protocol

All

FEATURES: Keyword Search Tabs

- Supports **multiple keywords** search separated by semicolons (;), representing **OR** logic.
- Combines all keyword tabs using **AND** logic for refined results.

Guide: A semicolon (;) should be used to separate keyword searches. This symbol represents OR in searching.

Example: Kaplan; Meier; Demographics; KM; Kaplan-Meier

The screenshot displays a search interface with three input fields labeled "Search Keyword Tab 1", "Search Keyword Tab 2", and "Search Keyword Tab 3". Each field is a white rectangle with a blue header and a small cursor icon at the bottom right. Between the first and second tabs, and between the second and third tabs, there is a blue button with the text "AND" in white. The entire interface is set against a light green background.

FEATURES: Exclusion Search Tabs

Users can **exclude unwanted keywords** to narrow results further.

Do you want to exclude some keywords in the filter?

Yes No

Exclude Keyword Tab 1

AND

Exclude Keyword Tab 2





FEATURES: Run Date Filtering

A toggle allows **filtering by run dates**, with a **date range selector** for precise control.

The screenshot shows a user interface for filtering by run dates. It features a question "Do you want to filter by run dates?" with two radio button options: "Yes" (selected) and "No". Below this is a dark blue button labeled "Select Date Range". Underneath the button is a date range selector with two input fields containing the dates "2003-01-13" and "2025-10-01", separated by a "to" label.

FEATURES: Footnotes Search

Introduced in version 1.3 and enhanced in version 1.4, **footnotes can now be searched and viewed** in the Details tab

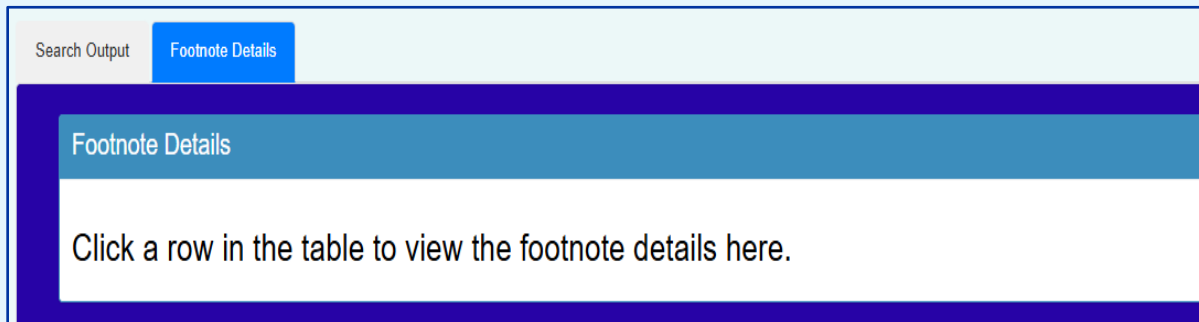


Table 14.3.1.2.3.1

Footnote: 8 a. If the same subject in a given treatment had more than one occurrence in the same preferred term event category, only the most severe occurrence is counted. If the same subject in a given treatment had more than one occurrence in the same preferred term event category, only the most severe occurrence is counted. Subjects are counted only once per treatment per event. For the TESS algorithm any missing severities have been imputed as severe unless the subject experienced another occurrence of the same event in a given treatment for which severity was recorded. In this case, the reported severity is summarized. Missing baseline severities are imputed as mild. Maximum Severity at any dictionary level is calculated after the report subset criteria is applied Includes data up to lag days after last dose of study drug. Percentages for gender specific events, use the corresponding gender count as a denominator. (@@) Denotes AE terms that were not coded by the MedDRA dictionary. &G_DICTVER coding dictionary applied.

OK

FEATURES: Direct Access Links

Total Search Rows: **270745**
 Results are for informational purposes only. Consult the asset statistician before use.

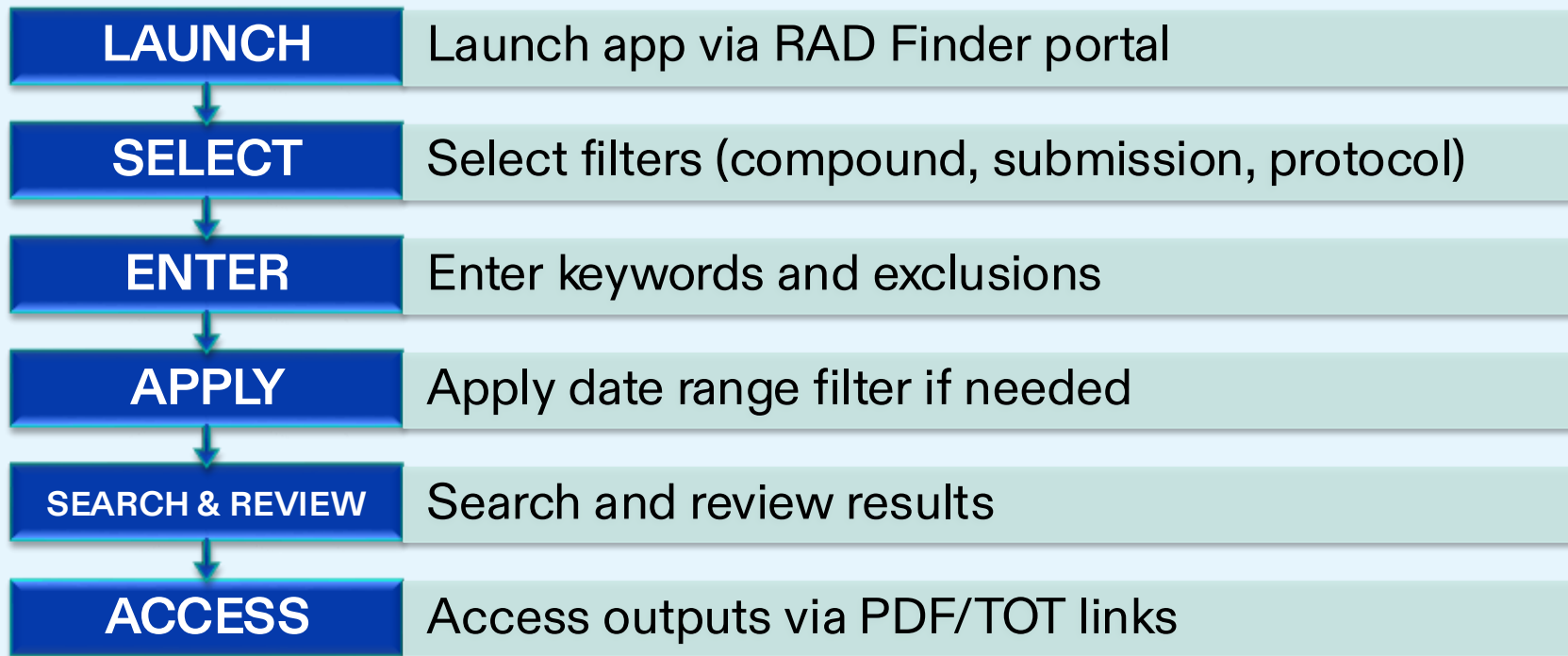
Search Output Footnote Details

Column visibility Download Show 20 entries Search:

TABLE NUMBER	TABLE TITLE	CDARS PDF LINK	CDARS PROGRAM LINK	FILE NAME	DATE GENERATED IN CDARS	FOOTNOTE
All	All	All	All	All	All	All
1	Table 13.7.21 Cumulative Number and Percent of Subjects with Antibody Usage for Treating BPAR (Full Analysis Set)	PDF	CDARS	abdrq_freq_fas_sum	2006 APR 10	View Footnote
2	Table 13.6.7.7 Cumulative Percent Of Subjects with One or More Acute Rejection Episodes Diagnosed Locally By Study Sites And Recorded As Adverse Events up to Month 6	PDF	CDARS	acutrej_freq_fas_sum	2006 APR 08	View Footnote
3	Table 0.1 MERGED PDF for A3921009 Month 3 Interim Analysis Tables	PDF	CDARS	adhctoc	2006 APR 08	View Footnote

- **PDF Link:** Opens the TLF file.
- **TOT Link:** Redirects to the TOT filename.

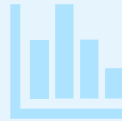
User Workflow



Benefits & Impact



**Reduces
retrieval
time from
hours to
minutes**



**Minimizes
redundant
analyses**



**Supports
faster
decision-
making**



**User-
friendly
interface for
non-
technical
users**

Limitations

**Metadata
completeness
dependency**

**Outputs for
informational use
only**

**Scope limited to
finalized TLFs in
repository**



Future Enhancements



Integration with
additional
repositories



Advanced
analytics and
visualization



AI-powered
semantic search



Continuous
improvement via
user feedback

Conclusion

RAD Finder transforms TLF retrieval

Aligns with digital transformation trends

Supports efficiency, transparency, and collaboration

Foundation for next-gen data ecosystems

References

- <https://www.forbes.com/councils/forbestechcouncil/2025/06/25/increasing-productivity-in-pharma-rd-starts-with-data/>
- Shiny for R (Posit) - <https://shiny.posit.co/>
- htmlwidgets for R - <https://www.htmlwidgets.org/>
- Mastering Shiny - <https://mastering-shiny.org/index.html>
- Pfizer Programming Master Handbook

Contact Information

Cielo Marie A. Castañeda

Email: cielomarie.castaneda@pfizer.com

LinkedIn: <https://www.linkedin.com/in/cielomariecastaneda/>

Acknowledgement

- ❖ Margaret Gamalo
- ❖ Gary Chen
- ❖ Kenneth Kwok
- ❖ Darrell Edgley
- ❖ Clifford Aspinall
- ❖ Christian Russel Reyes
- ❖ Mary Jane Cadatal
- ❖ Shixue (Melody) Liu
- ❖ Alexis Bernard Dalam
- ❖ Junibel De Leon



**The Global Healthcare
Data Science Community**

Contact Channels

- 📞 UK +44 1843 609600
- ✉ office@phuse.global
- 🌐 phuse.global

Social Media

- 🐦 [@phusetwitta](https://twitter.com/phusetwitta)
- 📘 [/phusebook](https://www.facebook.com/phusebook)
- ▶ [/phusetube](https://www.youtube.com/channel/UCphusetube)
- 🌐 [/company/phuse](https://www.linkedin.com/company/phuse)