Real World Evidence Community Forum

Capturing the Patient's Voice Through Lived Experience across Diaries & Social Media

phuse.global

Hosts & Presenters

connect·share·advance

Presenters:

Rachel Thackray & Francesca Toffolo (Adelphi Values);
Victoria Ikoro (Thermo Fisher Scientific)

Hosts:

- Elizabeth Merrall, *Johnson & Johnson;* Mary Anne Rutkowski, *Merck;* Nicky Newton, *PHUSE*

Content

connect·share·advance

- Introduction – 5’
- Rachel and Francesca– 20’
- Q&A – 5’
- Victoria – 20’
- Q&A – 10’

Please enter any questions you have for the presenters in the Q+A box within Zoom

RWE Working Group

connect·share·advance

https://phuse.global/Working_Groups

The Real-World Evidence Working Group aims to support, address and answer pertinent questions around real-world evidence. The Working Group is dedicated to sharing across the PHUSE Community (through Community Forums) and aligning on the best industry practices. Some of the questions we intend to address are:

- What are the requirements, technologies and processes/best practices needed to use RWD as a source for data analysis?
- incorporate real-world data into clinical trials?
- to support the use of real-world evidence as part of regulatory submissions?

Several Projects

- Applying Advanced Data Privacy Methods to Real World Data (RWD)
- Using OMOP and Other Real World Data Standards to Support Regulatory Submissions
- Estimands For RWD/RWE
- Quality & Reusability of RWD
- RWD Guideline for Programming & Analysis Processes
- Submitting RWD
- Missing Data Imputation in RWD Exploration of Multiple Techniques on Open-Source Data

Blog Articles & White Papers
Community forums
Conference activities

Estimand Webinar 21 May

US Connect Blog

If you would like to get involved, please email us:
workinggroups@phuse.global

Effective Use of Patient-Reported Daily Diary Data in Clinical and Psychometric Analysis: Balancing Benefits and Challenges for Optimal Outcomes

Rachel Thackray, Adelphi Values PCO, Bollington, UK
Francesca Toffolo, Adelphi Values PCO, Bollington, UK

Patient-Reported Daily Diaries

Patient Recorded data has Real-World Impact

Patient-reported daily diary data involves patients recording their symptoms and experiences regularly, often in real or near-real time.

These data are especially valuable in patient-centred outcome (PCO) trials used for psychometric analysis, helping to assess real-world treatment effects and inform clinical and regulatory decisions.

Electronic Recording

The use of electronic diaries and smartphone apps has made collecting daily diary data easier and more comprehensive.

Data Collection Challenges

Collecting and analysing daily diary data presents challenges, such as frequent data entries and potential for missing data and complex summaries.

Careful methodology and standardisation are needed to ensure daily diary data are reliable and interpretable for research.

What are Patient Reported Outcomes (PROs)?

Self-Assessments by Patients

Patient-reported outcomes (PROs) are self-assessments by patients about their own health status, symptoms, and everyday functioning, regarding the impact of disease and treatment.

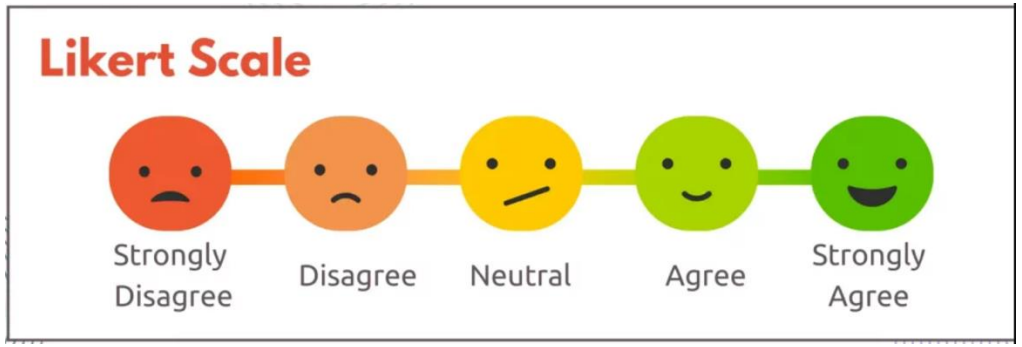
PROs provide important information that might not be readily captured by clinical or laboratory assessments alone.

Frequency and severity of symptoms, level of pain experienced, physical functioning, and emotional well-being may all be measured using PRO assessments.

What are Patient Reported Outcomes (PROs)?

PRO Data Collection

Structured assessment measures, known as instruments, are used to collect PRO data, which typically consists of individual items (questions) which result in a numeric response which fits to a predefined scale.



Activity	No difficulty	A little difficulty	Some difficulty	A lot of difficulty	Not able to do it
1. Lying in bed	1	2	3	4	5
2. Sitting	1	2	3	4	5
3. Getting in or out of bed or chair	1	2	3	4	5
4. Reaching or stretching	1	2	3	4	5

Related items may be grouped into domains (e.g., pain, physical function, emotional well-being), and sometimes all domains are combined into an overall score.

Daily Diary Data Collection

Mode of Data Collection

Patient-reported daily diary data is usually collected using **electronic devices** (e.g., tablets or smartphone apps), though paper diaries may be used in certain study populations or environments.

Electronic diaries can send automated reminders to help ensure timely completion of entries.

Scheduling

Diary entries can be planned for **fixed times of the day**, often with an **allowed time window** (e.g., ±2 hours).

Some studies require **event-based entries**, such as recording data during and after a medical incident (e.g., migraine), at specific time intervals.

Accuracy Measures

Lockout features can prevent entries outside allowed time windows, thus improving data quality.

Time stamps verifying when entries are made, may also support data accuracy.

Benefits and Challenges of Daily Diary Data

Benefits	Challenges
More **detailed and comprehensive information** than traditional, less frequent clinic-based assessments.	**Sustained patient engagement** is necessary and there is potential for reporting fatigue due to the repetitive nature of daily entries.
Real-time or near-real-time data collection **minimises recall bias** and offers a more accurate picture of symptom fluctuations and treatment effects.	High-frequency reporting can result in more **missing data** compared to less frequent clinic-based assessments.
Daily entries **capture intermittent symptoms**, day-to-day or even hour-to-hour variability, and events that may be missed during clinic visits.	This **missing data may introduce bias**, especially if linked to the patient's health status, complicating analysis.
Granularity leads to **richer datasets**, supporting more nuanced analyses and better clinical decision-making.	Effective study **design**, **patient support**, and **robust analytics** are crucial to maximise the benefits and minimise the drawbacks of daily diary data in research.
Data collection often occurs in the patient's home or usual environment, improving **patient comfort**. Completing entries in a familiar setting removes many environmental pressures and reduces biases related to clinical settings.	
The real-time approach **reduces the likelihood of patients forgetting details** or misremembering symptoms.	

Comparing Daily Diary Data to Another Instrument

Map Daily Diary Data to Scheduled Visits

Adelphi
ADELPHI VALUES

Patient-Centered Outcomes

Comparing Daily Diary Data to Another Instrument

Patient Global Impression of Severity (PGI-S) scale data are compared to a daily diary data measure

PGI-S scale ⟶ weeks 2, 4 and 8

Daily diary measure ⟶ study day 1 to 60.

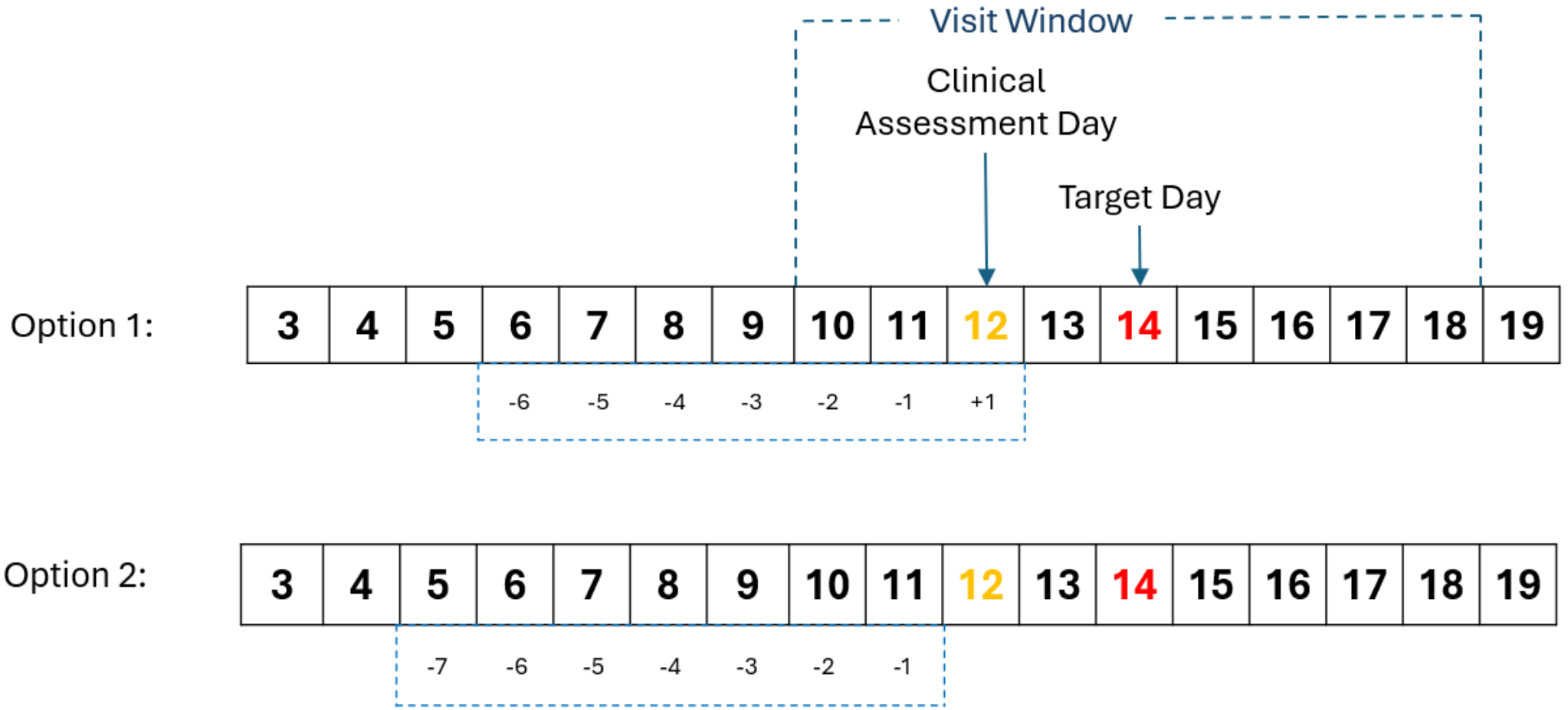
Windows around each visit are set to ±4 days

Comparing Daily Diary Data to Another Instrument

Aim: to compare a week's worth (seven days) of diary entries with the PGI-S scores obtained at each of weeks 2, 4, and 8.

Let's look at an example featuring the week 2 assessment.....

Week 2 Assessment Schedule



Inclusion of a Second Measuring Instrument

The study has decided that option 1 is the preferred method for defining the relevant 7 days for the assessment – days -6 to +1 from the day of the clinical assessment of PGI-S.

There is a **second instrument** used within this study: EQ-5D-5L scale ⟶ weeks 4 and 8.

The assessment date for EQ-5D-5L at week 4 differs from the date on which the PGI-S was assessed at week 4.

Let's look at some possible options.....

Inclusion of a Second Measuring Instrument

Two possible options:

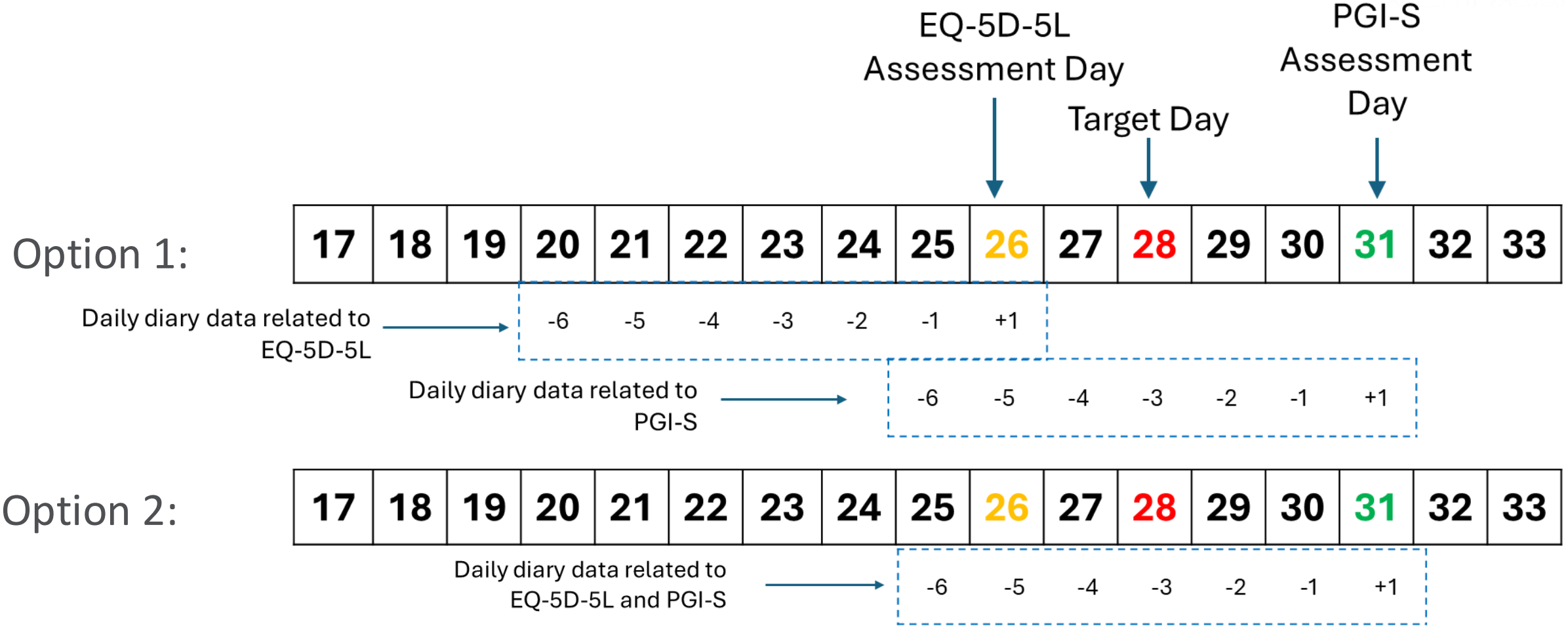
- Use a different set of 7 days for each instrument (based on each instrument's assessment date)

or

- Standardise the selection of diary days based on one instrument (e.g. PGI-S) regardless of which instrument is being compared.

There could be other options as well, these are only two examples!!

Week 4 Assessment Schedule



Deriving Weekly Score for Comparison with other Pre-Defined Anchors

ADELPHI VALUES

Patient-Centered Outcomes

Deriving a Weekly Score

Example

PRO instrument collected as a daily diary, with the following rules:

- Five questions (items) - responses collected daily.
- Item scores are integers ranging from 0 to 4.
- Missing data are null and not scored as 0.

Deriving a Weekly Score

There are (AT LEAST) two methods around deriving combined scores:

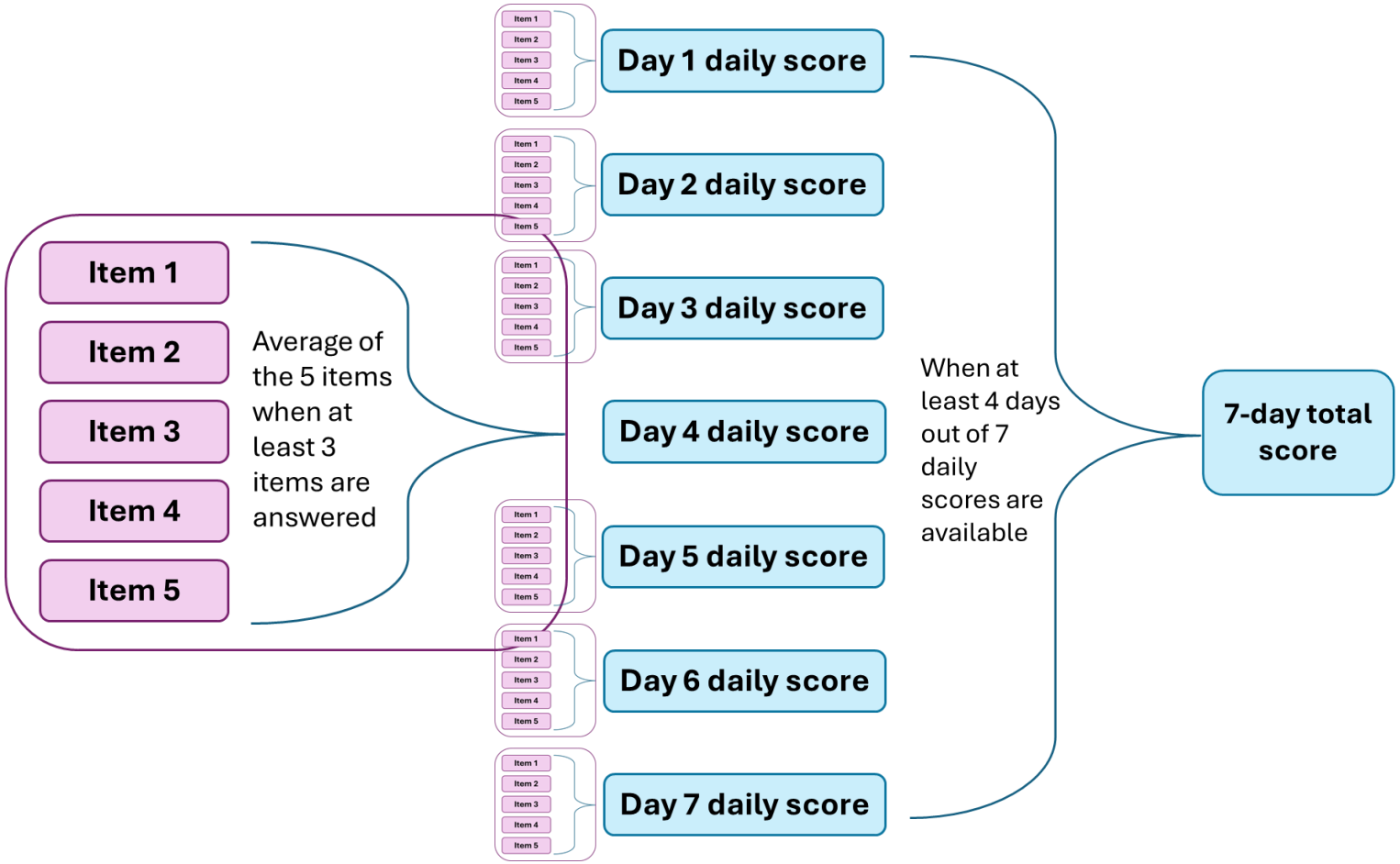
- **Method 1:** For a given day, **the daily score** is defined as the average of the 5 individual items score, when **at least 3 items** have a valid answer, otherwise it is missing (N/A).

- **Method 2: Item-level weekly score rule**: the item mean across 7 days is only computed if that item has responses on 4 or more days of the week (**4 out of 7**), otherwise that item's weekly mean is set to missing (N/A).

Let's look at 2 different methods to derive the weekly score.....

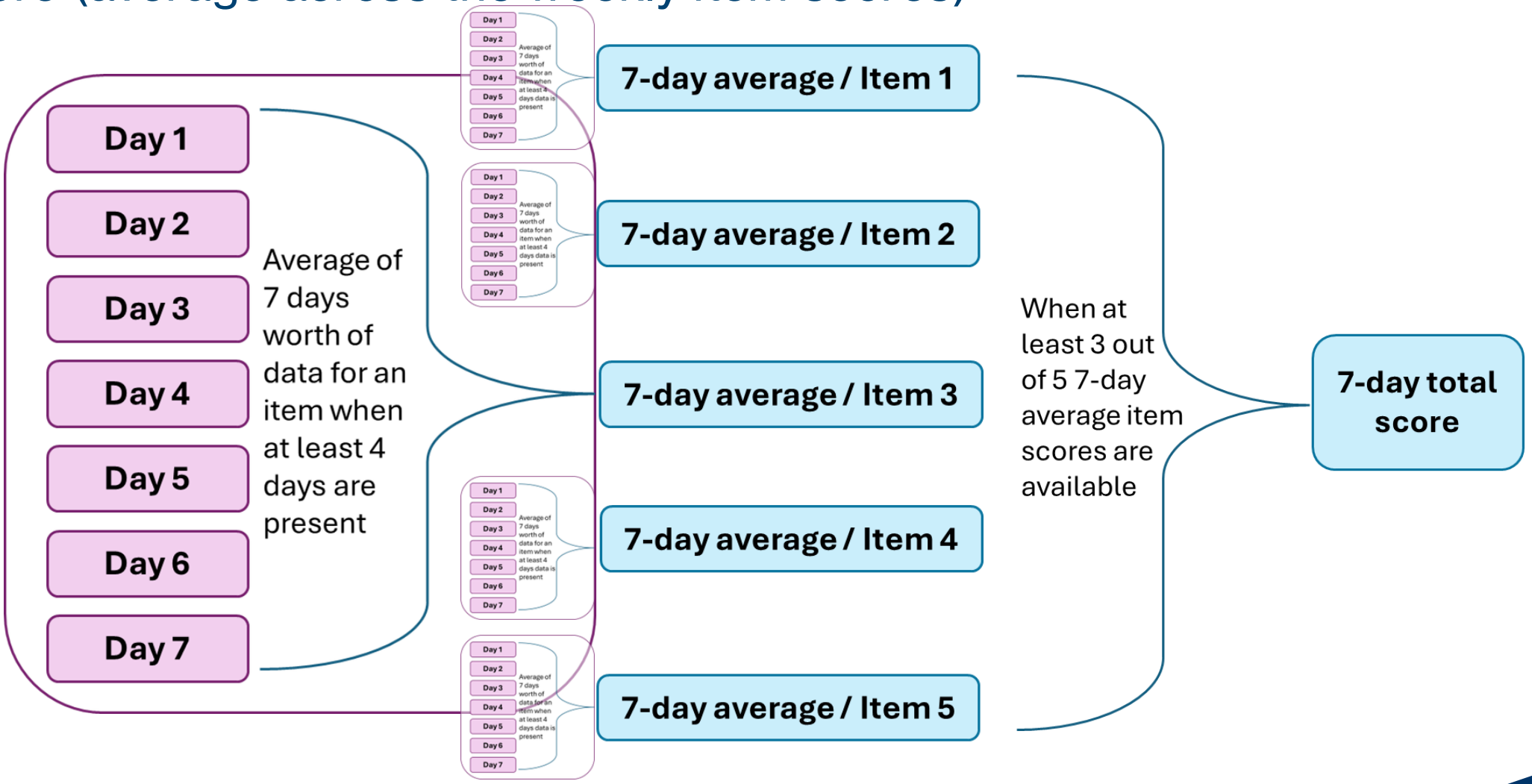
Method 1:

Total daily score (average across each day) ⟶ total weekly score (average across the daily scores)



Method 2:

Items weekly score (average across each item for the week) ⟶ total weekly score (average across the weekly item scores)



Deriving a Weekly Score – Example Data

Worked Example – data collected for all 5 items, from Day 1 to Day 7

	Day 1	Day 2	Day 3	Day 4	Day 5	Day 6	Day 7
Item 1	3	4	2				
Item 2	4	3	1	3	2	1	4
Item 3	0			3			
Item 4	3		0	2	2	1	
Item 5	3		4			3	

Method 1:

Total daily score (average across each day) ⟶ total weekly score (average across the daily scores)

	Day 1	Day 2
Item 1	3	4
Item 2	4	3
Item 3	0	
Item 4	3	
Item 5	3	

- A total daily score is derived for each day (if 3/5 scores present).
- e.g. Day 1 is calculated as: (3+4+0+3+3) / 5 = 2.6.
- Day 2 cannot be calculated because only 2 items have been answered.
- Total daily scores shown below.
- 7-day Weekly Score is derived if 4/7 daily scores are present.

	Day 1	Day 2	Day 3	Day 4	Day 5	Day 6	Day 7	7-day score
Total Score	2.6	NA	1.8	2.7	NA	1.7	NA	2.2

Method 2:

Items weekly score (average across each item for the week) ⟶ total weekly score (average across the weekly item scores)

	Day 1	Day 2	Day 3	Day 4	Day 5	Day 6	Day 7
Item 1	3	4	2				
Item 2	4	3	1	3	2	1	4
Item 3	0			3			
Item 4	3		0	2	2	1	
Item 5	3		4			3	

- Each item's 7-day average score is derived (if 4/7 scores present).
- Item 1 can't be derived as only 3 scores present.
- Item 4 can be derived (3+0+2+2+1)/5= 1.6.
- 7-day Weekly Score is derived if 3/5 7-day average scores are present.
- **Weekly score cannot be derived using this method.**

	Can average be calculated?	7-day average score
Item 1	No, less than 4 days collected	N/A
Item 2	Yes	2.6
Item 3	No, less than 4 days collected	N/A
Item 4	Yes	1.6
Item 5	No, less than 4 days collected	N/A

Weekly Score Summary

- Weekly total score for diary data can be derived in different ways; in this presentation we illustrated 2 methods:
  - Method 1 (deriving a total daily score first) weights days with enough data over items.
  - Method 2 (deriving a weekly item score first) weights items with enough repeated measures over days.
- Neither method is wholly correct or incorrect, they answer slightly different questions and add weight differently.
- Generally, this is only an issue with very sparse data but should be considered if there is a lot of missing data.
- Rules on which 7-days to use to map to scheduled visits and on derivation of weekly scores should be pre-specified in the analysis plan prior to data analysis.

Conclusion

- Patient-reported daily diary data offer **unique and invaluable insights into patients' real-world experiences**, allowing the capture of symptom fluctuations and treatment effects that are often missed by traditional clinic-based assessments.

- To maximise the value of daily diary data, **methodologies must be carefully pre-specified** to ensure suitable instruments are used at the correct frequency to support patient engagement, and **utilise electronic solutions** that facilitate timely, accurate, and complete reporting.

- When applied effectively, daily diary data can **enhance the reliability, relevance and completeness of patient-centred data** within clinical research, supporting better decision-making for patients, clinicians, and regulators alike.

Questions?

Rachel Thackray, Francesca Toffolo
Adelphi Values PCO
Email: Rachel.Thackray@omc.com
Francesca.Toffolo@omc.com

Boston, USA
T: +1 617 720 0001

Manchester, UK
T: +44 (0)1625 576815

Q&A

Translating the Patient Voice from Social Media into PRO-Aligned Domains Using NLP

Victoria Ikoro

Senior Data Scientist, Thermofisher Scientific

The world leader in serving science

Agenda

1 **Background**

Capturing patient experience beyond clinical data

2 **Methods**

Translating social media narratives using NLP

3 **Results**

Identifying key burden domains

4 **Interpretation**

Contextualizing findings with PRO frameworks (QOLIE-31)

5 **Implications**

Informing patient-centered measurement

Traditional RWD Captures Clinical Events — Not Lived Experience

Traditional real-world data (EHR / claims) captures:

- Diagnoses and disease classification
- Treatment exposure and medication us
- Healthcare utilization (visits, hospitalizations)

However, it often lacks visibility into:

- Emotional burden (fear, anxiety, distress)
- Cognitive challenges (memory, focus, brain fog)
- Social impact (independence, work, driving)
- Day-to-day patient experience

Patients Share Their Lived Experience Outside Clinical Settings

- **Patients frequently describe their experiences in:**
  - Online forums
  - Social media communities
- **These narratives include:**
  - Seizure impact on daily life
  - Emotional distress and uncertainty
  - Cognitive and functional challenges
- **Provides unfiltered, real-world patient perspective**

Bottom Takeaway

Social media offers a rich, underutilized source of patient-reported experience

Data Sources and Characteristics

Online epilepsy communities provide large-scale patient narratives (2018–2025)

Data sources

- Coping With Epilepsy (coping-with-epilepsy.com)
- Epilepsy Foundation Community Forums (epilepsy.com)
- EpilepsyDisease.com (epilepsydisease.com)
- Reddit – r/epilepsy (reddit.com/r/Epilepsy/)

Dataset scale

- ~98,000 posts collected
- ~8,200 users
- ~21,800 posts containing relevant search terms

Data characteristics

- Publicly available online discussions
- Real-world, unfiltered patient narratives
- Topics include seizures, treatment, and daily life impact

Identifying Relevant Epilepsy Discussions

A targeted set of keywords was developed to identify drug-resistant epilepsy (DRE) discussions, including patients with or without neuromodulation

Terms were informed by:

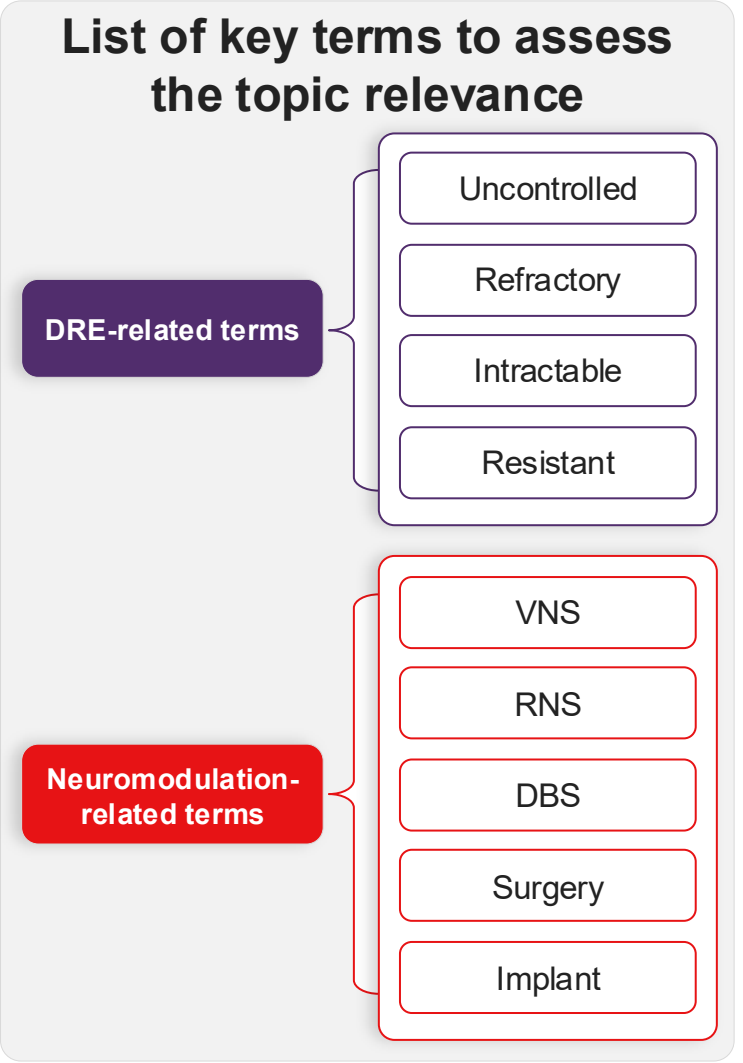
- Synonyms for DRE (e.g., refractory, intractable, uncontrolled)
- Neuromodulation terminology (e.g., VNS, RNS, DBS, surgery, implant)
- Initial exploration of language used in patient discussions

Search Strategy:

- Neuromodulation terms were captured using:
  - Acronyms (VNS, RNS, DBS)
  - Full terminology (vagus nerve stimulation, responsive neurostimulation, deep brain stimulation)
- Posts were subset to include those containing any relevant keyword

Outcome:

- Focused identification of high-burden epilepsy discussions, as DRE populations represent patients with greater disease burden and unmet need
- Reduced noise from unrelated content



Isolating Patient Voice Using LLM-Based Classification

- Distinguish patient-authored posts from other perspectives in social media data
- Ensure analysis reflects true patient-reported experience

Stage	Number of Posts	Number of Users
Raw dataset	98,000	8,000
After keyword filtering (DRE / neuromodulation)	21,806	6,756
After perspective classification	14,363	5,073
Final patient dataset	**14,363**	

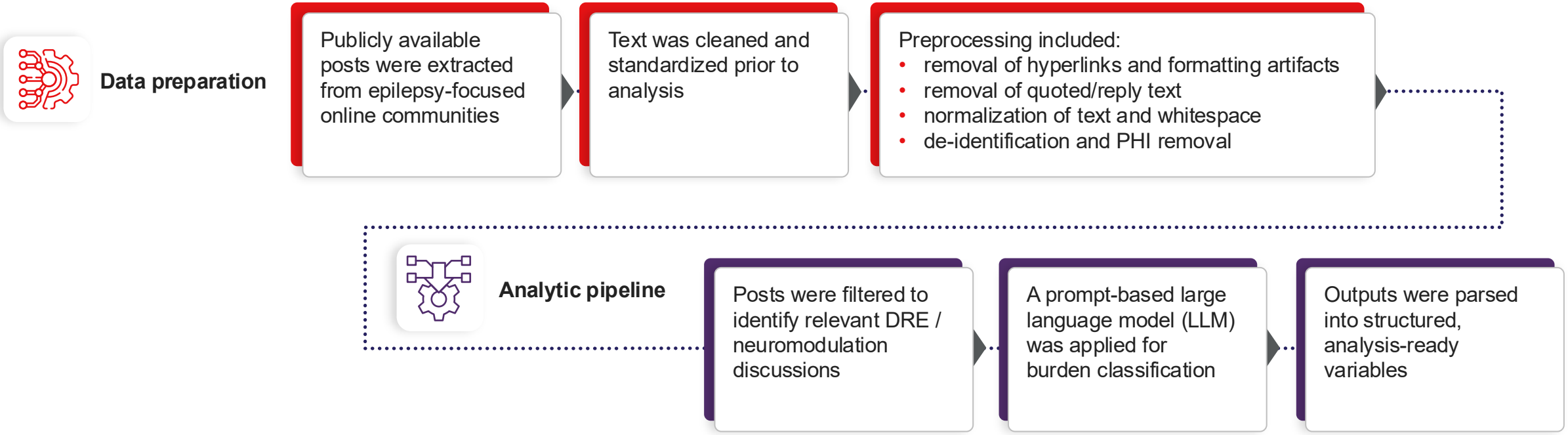
Approach:

- A large language model (LLM) was used to classify each post into:
  - **Patient** (first-person lived experience)
  - **Caregiver** (describing another individual's experience)
  - **Other** (general discussion or unclear perspective)

Implementation:

- Classification based on linguistic patterns and contextual cues
- Guardrails applied to identify caregiver language and reduce misclassification
- Dataset refined to include patient-only narratives
- Removed caregiver and non-experiential content
- Ensures burden estimates reflect patient-reported experience rather than proxy perspectives

Text Preparation and Analytic Pipeline



LLM-Based Classification of Patient Narratives into Burden Domains

Approach

Patient posts were analyzed using a prompt-based large language model (LLM)

A structured prompt mapped each post to predefined burden domains and functional markers

Domains were defined a priori based on clinically relevant constructs and established PRO frameworks (e.g., QOLIE-31)

Burden Domains Captured


Treatment burden (medication effects, treatment complexity)


Emotional burden (anxiety, fear, distress)


Seizure unpredictability (uncertainty, lack of control)


Cognitive burden (memory, concentration difficulties)


Work/daily impact (employment, daily functioning)

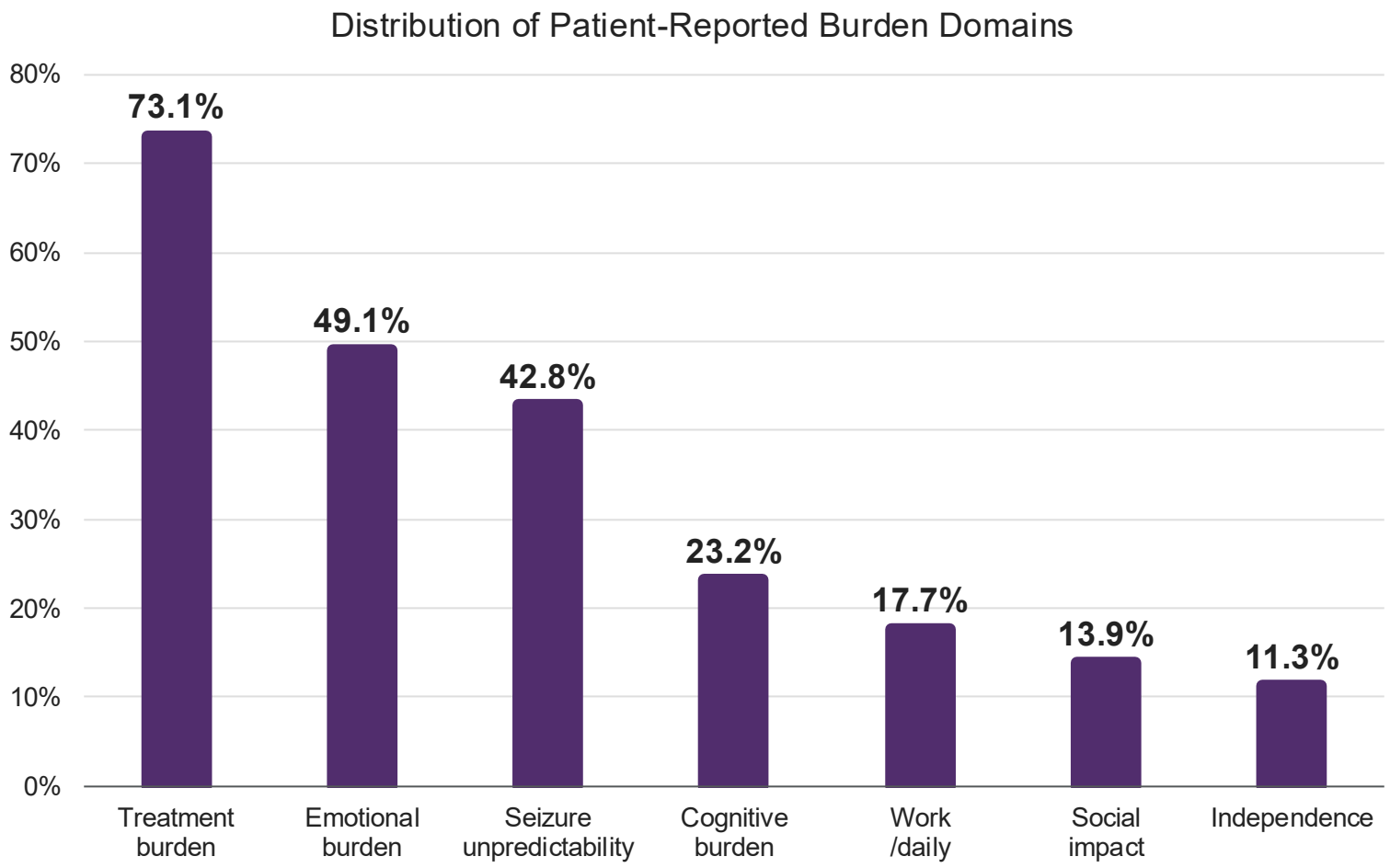

Social impact (relationships, social participation)


Independence (driving, autonomy)

- Each burden domain was assigned as a binary indicator: **1 = present, 0 = absent**
- Multi-label classification allowed assignment of multiple burden domains per post
- Functional markers (e.g., driving restriction, employment impact) were also extracted
- A subset of classifications was manually reviewed to assess consistency and validity of LLM-derived labels."

LLM-based classification transforms unstructured patient narratives into structured, multi-domain variables aligned with patient-centered outcome frameworks. Structured outputs enabled quantitative analysis of domain frequency and co-occurrence

Treatment Burden Dominates Patient-Reported Experience

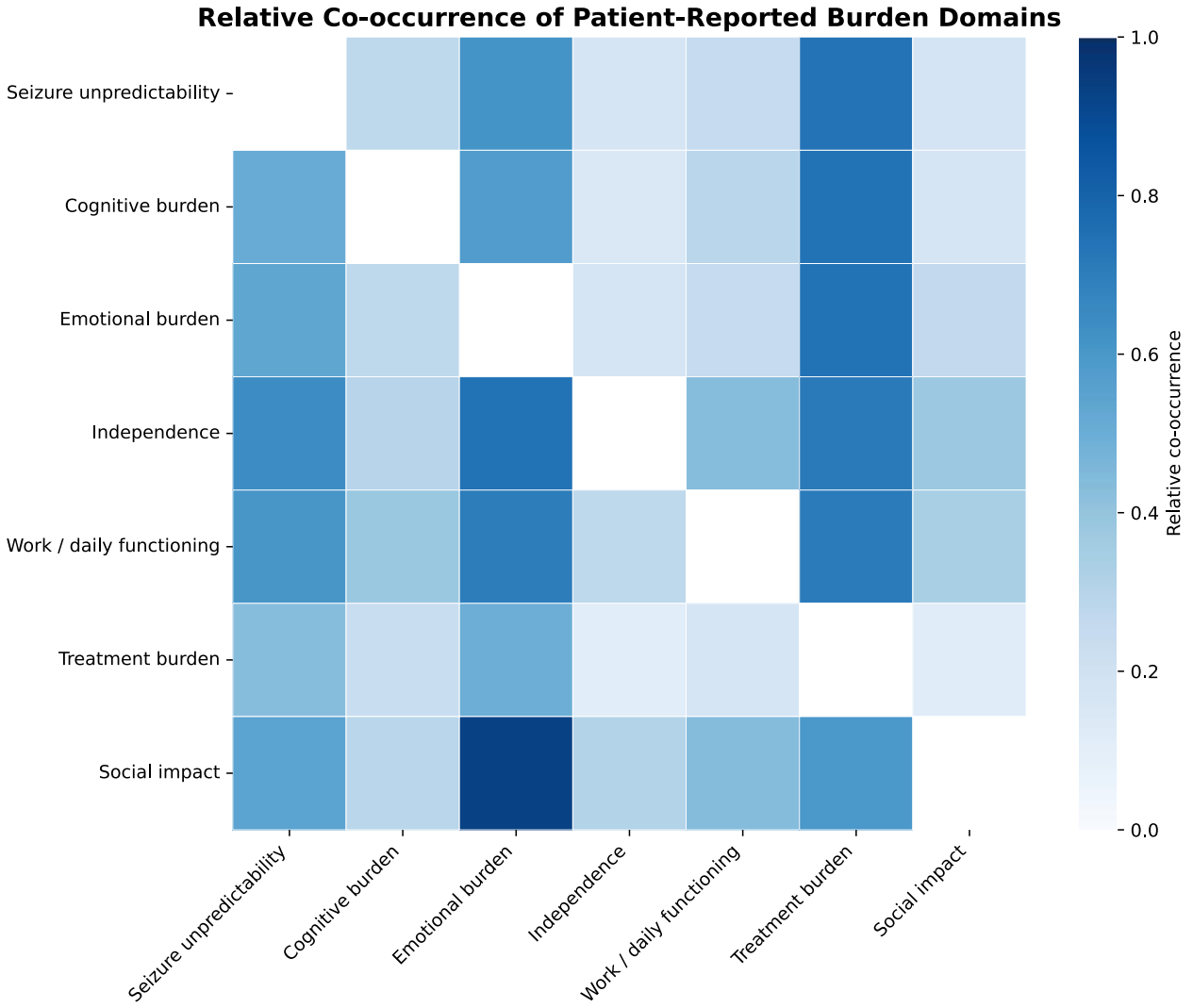


Key Observations:

- Treatment burden is present in the highest proportion of posts (~73%)
- Emotional burden (~49%) and seizure unpredictability (~43%) are also frequently reported
- Cognitive and functional impacts occur less frequently but remain notable

Percentages represent the proportion of posts in which each domain is present; domains are not mutually exclusive.

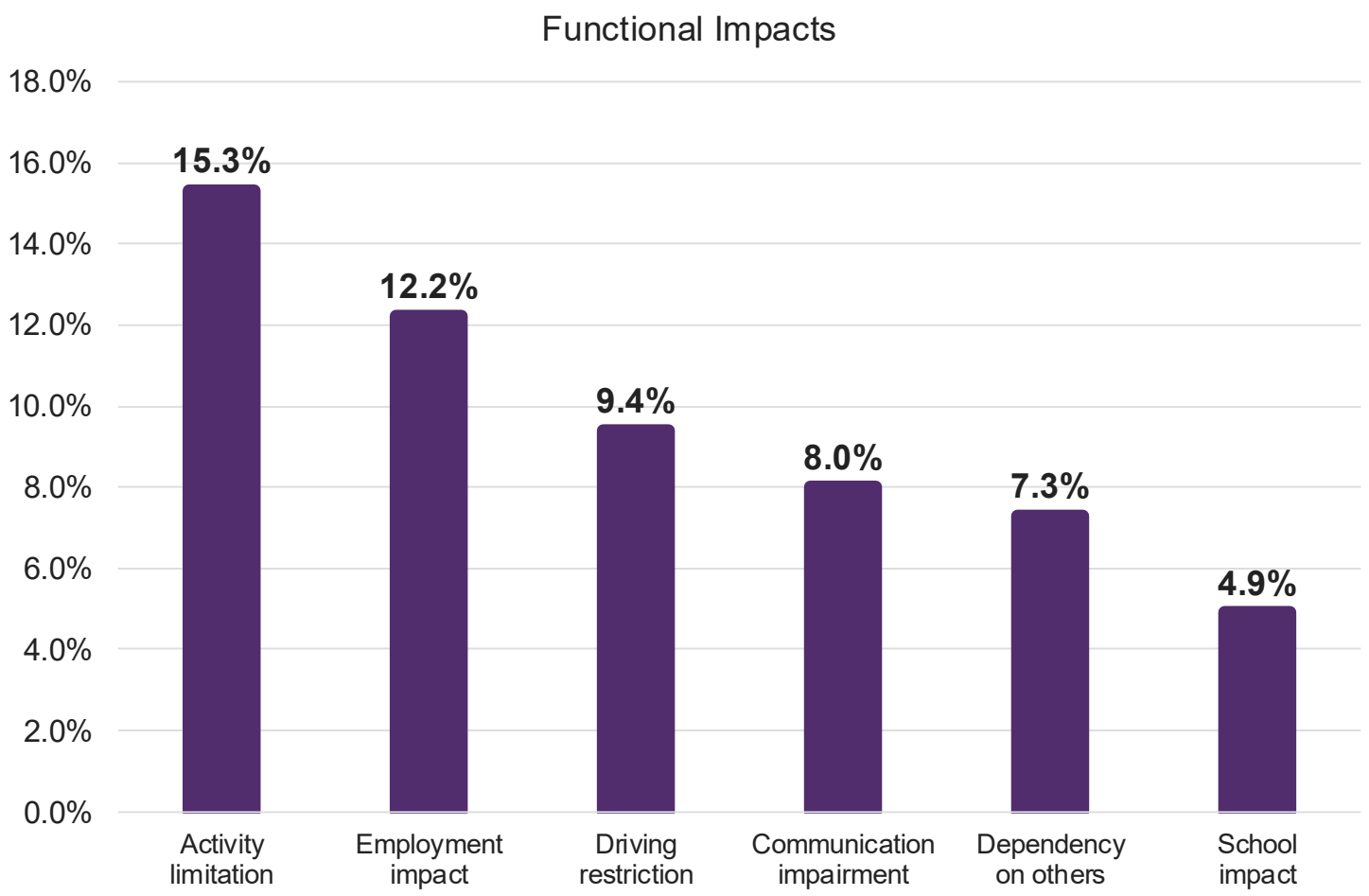
Epilepsy Burden is Multi-Dimensional and Interconnected



- **Multiple burden domains frequently co-occur within individual posts**
  - ~70% of posts include ≥2 burden domains, highlighting the multi-dimensional nature of patient burden.
- **Strong co-occurrence observed between:**
  - Emotional burden and social impact
  - Treatment burden and seizure unpredictability
- **Moderate overlap observed between:**
  - Emotional and cognitive burden
  - Work/daily functioning and treatment burden

Cells represent relative co-occurrence of burden domains; darker shading indicates stronger overlap. Domains are not mutually exclusive.

Daily Life Disruptions Reveal Real-World Consequences of Epilepsy



Key Observations

- Functional impacts are reported in a subset of posts
- Activity limitation and employment impacts are most frequently observed
- Driving restriction remains a notable impact on patient experience

While less frequent than core burden domains, functional impacts highlight meaningful real-world consequences of epilepsy

Treatment Burden Emerges as the Central Patient Concern

"I've been on so many medications over the years, and each one comes with side effects. Some made me feel worse than the seizures themselves. It's exhausting trying to find something that works."

"The medication helps, but the side effects are hard to deal with. I feel foggy, tired, and not like myself. Sometimes I wonder if it's worth it."

"Since this medication routine started, I haven't felt like myself. I feel depressed, confused, and overwhelmed most days. It's hard to deal with the constant changes."

Treatment burden present in 73.1% of posts■ Treatment burden is the most frequently reported domain across patient narratives
Patients commonly describe:
medication side effects
frequent treatment adjustments
incomplete seizure control

Treatment burden often co-occurs with emotional distress & cognitive challenges

42

Examples of Emotional Burden

"I'm always worried about going back to bed in case I have another seizure. When it happens, I feel disoriented and scared, and it takes time to recover. It still scares me every time."

"I feel anxious all the time not knowing when the next seizure will happen. The unpredictability makes it hard to relax or plan anything. It's exhausting to live like this."

"Since this all started, I haven't felt like myself. I feel depressed, confused, and overwhelmed most days. It's hard to deal with the constant changes."

Emotional burden present in **49.1% of posts**

Emotional burden is one of the most frequently reported domains

Patients frequently describe **anxiety, fear, and uncertainty** related to seizures

Emotional burden often co-occurs with treatment challenges and seizure unpredictability

Alignment with QOLIE-31 Reveals Gaps in Treatment Burden Measurement

QOLIE-31 Captures:

- Medication side effects
- Mental/cognitive effects
- Concern about long-term use

Patient Experience of Treatment (This Study):

- Ongoing trial-and-error with medications
- Continued seizures despite treatment
- Frequent treatment changes
- Managing a complex, evolving treatment journey

QOLIE conceptualizes treatment burden primarily as medication side effects. However, patients describe treatment as an ongoing process — involving repeated medication changes, continued seizures, and managing a complex treatment pathway

Patient-Reported Burden Reveals Multi-Dimensional Disease Impact Beyond Seizure Activity

Treatment burden is more frequently reported than seizure-related burden, highlighting the impact of therapy on daily life

Emotional burden is highly prevalent, indicating that patient experience extends beyond clinical symptoms

Functional impacts are present but less frequently reported, suggesting variability in how burden is expressed

- Burden domains frequently co-occur, demonstrating that patient experience is multi-dimensional
- Patient-reported experience differs from traditional clinical focus, emphasizing treatment challenges, emotional burden, and overlapping real-world impacts

Patient Voice Complements Traditional Real-World Data

Traditional RWD (EHR / Claims)

Diagnoses

Treatment exposure

Healthcare utilization

Structured data

Clinician-reported

+

Patient Voice (Social Media)

Lived experience

Treatment burden

Daily challenges

Unstructured narratives

Patient-reported

Implications for Patient-Centered Measurement and Research

For Measurement of Burden

- Treatment burden emerges as a major component of patient experience and should be explicitly captured
- Emotional burden is highly prevalent and should be considered a core domain
- Unstructured analysis can help design structured assessment design for a therapeutic area

For Study Design

- Patient experience is multi-dimensional, suggesting the need for multi-domain assessment
- Single endpoints may not fully capture the breadth of patient burden

For Real-World Evidence

- Patient voice provides insight into aspects of burden not captured in structured data
- Can complement traditional RWD to improve understanding of disease experience

Patient-reported data highlights important aspects of burden—particularly treatment and emotional impact—that may not be fully captured in traditional approaches

Limitations

- Social media data may not be fully representative of the broader epilepsy population
- Posts are self-reported and not clinically verified
- **However, these data capture real-world patient experiences and perspectives not observable in structured datasets**

Conclusion

- Social media narratives can be translated into structured, PRO-aligned domains using LLM-based methods
- Patient-reported burden is multi-dimensional, extending beyond clinical measures
- Treatment and emotional burden emerge as central components of patient experience
- Patient voice provides population-level insight into lived experience that is not captured in traditional data sources

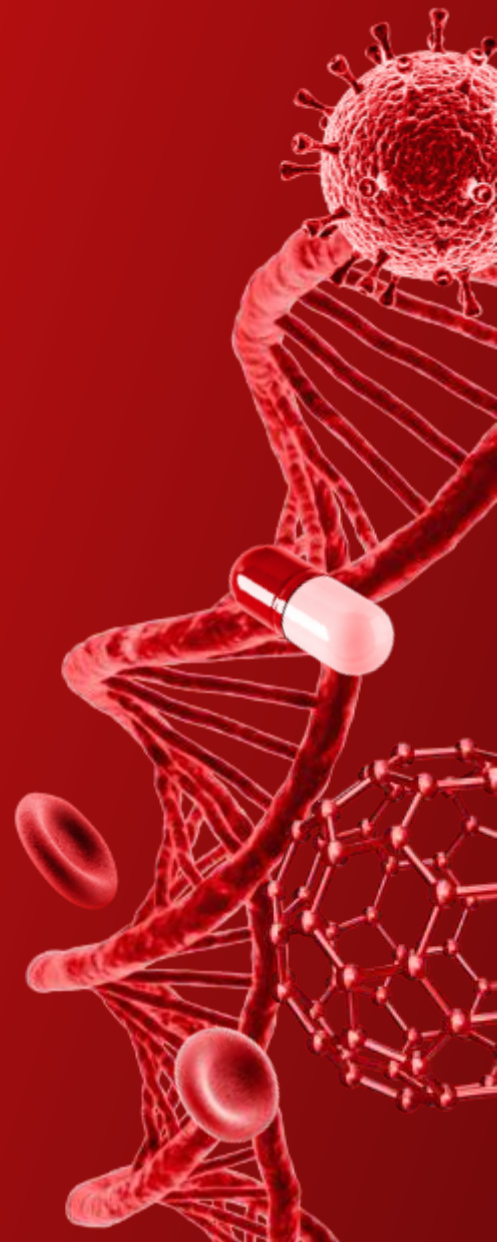
Patient voice analytics complements traditional real-world data by adding context on how disease and treatment are experienced, enabling a more comprehensive understanding of disease burden

ThermoFisher SCIENTIFIC

Questions?

The world leader in serving science

Thank you



Q&A

Keep in Touch!

connect·share·advance

If you would like to get involved, email: workinggroups@phuse.global

https://phuse.global/Working_Groups

Slides & recording will shortly be available on our event page

Welcome to reach out to any of the presenters directly:

- Rachel and Francesca: rachel.thackray@omc.com
  francesca.toffolo@omc.com
- Victoria: victoria.ikoro@thermofisher.com

Coming up in July

connect·share·advance

- **GenAI and its impact on RWD**
  - Presenters to include runners-up of our Catalyst Challenge
  - Expected date/time: Thursday 9 or 16 July – 9am ET/ 2pm BST/ 3pm CET
  - Registration to open in June, watch this space!

Real-World Evidence Virtual Event

- 30 September & 1 October
- 8 Selected High Quality Presentations
- Q & A sessions
- Working group updates

- Call for papers open NOW!

- Organizing Committee:
  - Working group leads
  - Community forum leads

connect·share·advance

THANK YOU!