

2025 Conference Poster Portfolio

PPD™ Evidera™ Solutions

PPD™ CorEvitas™ Solutions

Conferences

The Professional Society for Health Economics and Outcomes Research (ISPOR NA)	4
Biostatistics & Health Analytics	18
International Society for Pharmacoepidemiology (ISPE)	22
International Academy of Health Preference Research (IAHPR)	50
GetReal	56
ObesityWeek	80
The Professional Society for Health Economics and Outcomes Research (ISPOR EU)	84

How to navigate around the portfolio



Returns to the poster list



Returns to the conference list

Click conference title to take you to that section

Click poster title to take you to that poster

Conferences

The Professional Society for Health Economics and Outcomes Research (ISPOR NA)	4
Biostatistics & Health Analytics	14
International Society for Pharmacoepidemiology (ISPE)	18
International Academy of Health Preference Research (IAHPR)	46
GetReal	52
ObesityWeek	76
The Professional Society for Health Economics and Outcomes Research (ISPOR EU)	80

The Professional Society for Health Economics and Outcomes Research (ISPOR NA)

13-16 May 2025 | Montreal, Quebec, Canada

Posters

Mind the Gap: Delays between Vaccine Licence and NITAG Recommendation in the US vs. Europe
Exploring Treatment Preferences for Rare Diseases: A Systematic Review of Quantitative Preference Studies
Stability of a Large Language Model for Data Extraction in Systematic Literature Reviews
Leveraging Real-world Data vs. External Trials to Inform Comparative Efficacy within Single-arm Trial-based NICE Submissions

The Professional Society for Health Economics and Outcomes Research (ISPOR NA)

13–16 May 2025 | Montreal, Quebec, Canada

Posters

A Real-World Perspective of Quadruple Treatment Patterns for 1L Multiple Myeloma Patients Across North and South America, Europe, and Asia

Post-DCE Qualitative Interviews: A Novel Method to Understand the Rationale for Stated Preferences – An Example Among Patients and Physicians in Severe Asthma

Mind the Gap: Delays between Vaccine Licensure and NITAG Recommendation in the US vs. Europe

Exploring Treatment Preferences for Rare Diseases: A Systematic Review of Quantitative Preference Studies

Stability of a Large Language Model for Data Extraction in Systematic Literature Reviews

Leveraging Real-world Data vs. External Trials to Inform Comparative Efficacy within Single-arm Trial-based NICE Submissions



Poster Title

A Real-World Perspective of Quadruple Treatment Patterns for 1L Multiple Myeloma Patients Across North and South America, Europe, and Asia

Objective

- The objective of this study is to identify and examine QD treatment patterns among 1L MM patients across different countries in the top 10 markets (Brazil, Mexico, US, France, Germany, Italy, Spain, UK, China, and Japan).



HSD81

A Real-World Perspective of Quadruple Treatment Patterns for 1L Multiple Myeloma Patients Across North and South America, Europe, and Asia

Tatiana Checchia, MBA¹; Kelechi Emeanuru, MPH²; Mariah Baltezgar, MBA²; Luis Peloso, MD, PhD³
¹Thermo Fisher Scientific, Prague, Czech Republic; ²Thermo Fisher Scientific, Waltham, MA, USA; ³Thermo Fisher Scientific, Sao Paulo, Brazil

Background

- Recent combination therapies, such as quadruple (QD) therapy, have become the preferred first-line (1L) treatment option for multiple myeloma (MM) patients.
- These QD combinations have proved to be safer and more effective for MM patients, offering prolonged and sustained responses over time.
- Changes to the 2025 NCCN guidelines include an update to the list of primary regimens for MM 1L in stem cell transplantation (SCT) eligible patients, moving the QD therapy, daratumumab + bortezomib + lenalidomide + dexamethasone (D-VRd), from other recommended regimens to a preferred regimen (category 1) and adding the QD therapy, isatuximab-irfc + bortezomib + lenalidomide + dexamethasone (Isa-VRd) to other recommended regimens.
- Isa-VRd QD therapy was also added for MM 1L SCT ineligible patients as preferred regimens (category 1).

Objectives

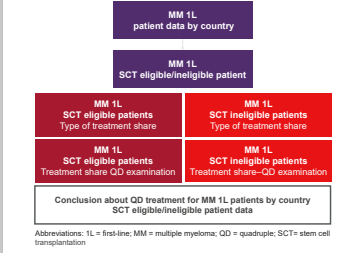
- The objective of this study is to identify and examine QD treatment patterns among 1L MM patients across different countries in the top 10 markets (Brazil, Mexico, US, France, Germany, Italy, Spain, UK, China, and Japan).

Methods

Patients who received at least one line of therapy were identified through PPD[™] Oncolocator[™] Global Cancer Treatment Patterns[™], a real-time data collection tool, to examine treatment patterns reported by healthcare providers (HCPs) (Figure 1).

Data were collected on 1L MM patients from December 2022 to November 2024 who were treated with a drug combination that specifically includes a proteasome inhibitor, an immunomodulatory drug, a steroid, and anti-CD38 called as QD.

Figure 1. PPD[™] Oncolocator[™] Global Cancer Treatment Patterns MM 1L Data Examination



Abbreviations: 1L = first-line; MM = multiple myeloma; QD = quadruple; SCT = stem cell transplantation

Results

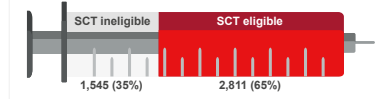
- A total of 4,356 patients were included in the cohort, with an average of 425 patients per country. Among those 1L MM patients, 65% (n=2,811) were SCT eligible, and 35% (n=1,545) were SCT ineligible (Table 1) and (Figure 2).

Table 1. MM 1L Patients by Country

Country	Total No. of MM 1L Patients	Total % of MM 1L Patients
Italy	551	13%
Spain	547	13%
Brazil	497	11%
China	478	11%
Japan	478	11%
US	464	11%
Germany	400	9%
France	388	9%
UK	333	8%
Mexico	220	5%
Grand Total	4,356	100%

Abbreviations: 1L = first-line; SCT = stem cell transplantation
 Date: PPD[™] Oncolocator[™] Global Cancer Treatment Patterns[™], November 2022–December 2024, MM 1L patients (n=4,356)

Figure 2. MM 1L Patients by SCT Status



Abbreviations: 1L = first-line; MM = multiple myeloma; SCT = stem cell transplantation
 Note: SCT eligible includes patients who have undergone SCT.
 Date: PPD[™] Oncolocator[™] Global Cancer Treatment Patterns[™], November 2022–December 2024, MM 1L patients (n=4,356)

MM 1L SCT Eligible Patients

- Among 1L SCT eligible patients (n=2,811), the average number of patients per country was 281 patients (n=131–434).
- QD treatment use was highest in France (89%), followed by Germany (82%), Italy (63%), and the UK (61%) (Figure 3).

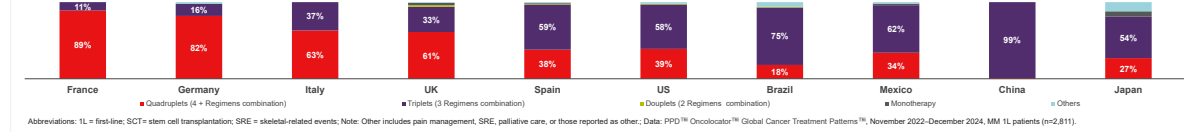
Conclusions

- Contemporary real-world data on QD trends reflect an evolving landscape, particularly with treatment patterns for SCT eligible MM patients across the US and EU regions.
- Examining QD standards and trends across various countries and time periods can help stakeholders and communities enhance their understanding of the adoption rates of the most recent treatment approvals.
- Monitoring data on QD patterns will provide valuable insights into any persistent market changes, as additional assessments are needed to evaluate the global use of QDs.

References

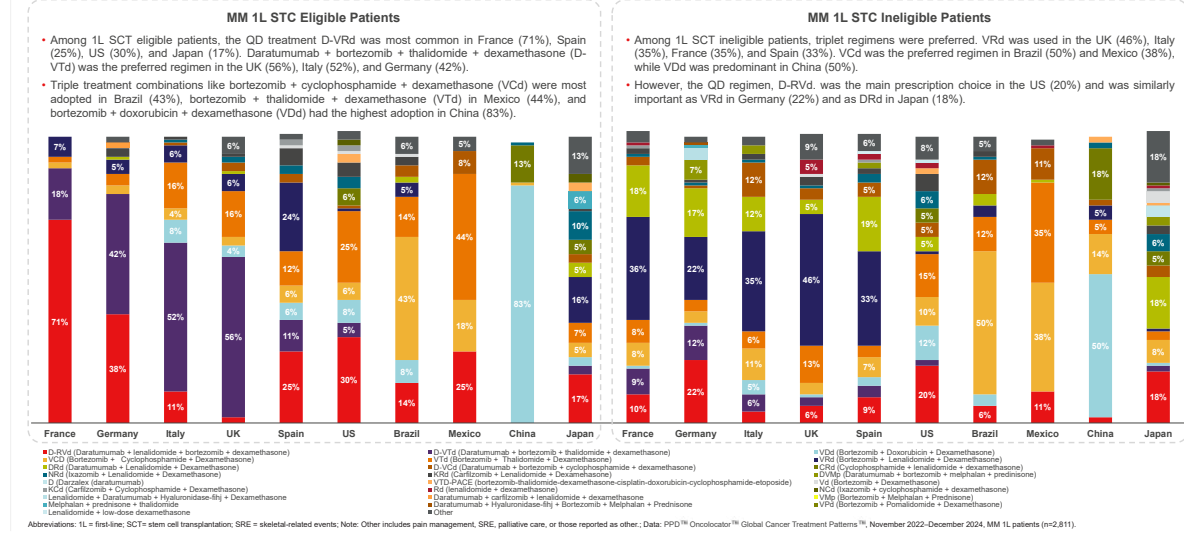
1. SEER. Cancer stat facts: myeloma. <https://seer.cancer.gov/statfacts/html/mymy.html>. 2. Siegel RL, et al. CA Cancer J Clin. 2023;73(1):17–48. 3. NCCN Guidelines. Multiple Myeloma. https://www.nccn.org/professionals/physician_glg/pdf/multiple_myeloma.pdf. 4. Dimopoulos MA, et al. Ann Oncol. 2021;32(3):309–322.

Figure 3. Treatment Share by Combination Type (MM 1L SCT Eligible Patients)



Abbreviations: 1L = first-line; SCT = stem cell transplantation; SRE = skeletal-related events; Note: Other includes pain management, SRE, palliative care, or those reported as other.; Data: PPD[™] Oncolocator[™] Global Cancer Treatment Patterns[™], November 2022–December 2024, MM 1L patients (n=2,811).

Figure 4. Treatment Share by Country (MM 1L SCT Eligible and Ineligible Patients)



Abbreviations: 1L = first-line; SCT = stem cell transplantation; SRE = skeletal-related events; Note: Other includes pain management, SRE, palliative care, or those reported as other.; Data: PPD[™] Oncolocator[™] Global Cancer Treatment Patterns[™], November 2022–December 2024, MM 1L patients (n=2,811).



Poster Title

Post-DCE Qualitative Interviews: A Novel Method to Understand the Rationale for Stated Preferences – An Example Among Patients and Physicians in Severe Asthma

Objective

- We present a novel qualitative post-DCE interview method helpful in elucidating individuals' rationale for stated preferences and deriving rich insights into the drivers of the heterogeneity often observed in stated preference studies.

AUTHORS

Heather Gelhorn¹
Hannah Collacott¹
Melissa Ross¹

AFFILIATION

¹Thermo Fisher Scientific, Waltham, MA, USA

PRACTICE AREA

PPD™ Evidera™ Patient-Centered Research, Thermo Fisher Scientific

CLIENT

INTERNAL

Post-DCE Qualitative Interviews: A Novel Method to Understand the Rationale for Stated Preferences – An Example Among Patients and Physicians in Severe Asthma

Heather Gelhorn, PhD¹; Hannah Collacott, MSc¹; Melissa Ross, PhD¹
¹Thermo Fisher Scientific, Waltham, MA, USA

Background

- Patient preference (PP) methodologies, such as discrete choice experiments (DCEs), help elucidate what is important to patients by quantifying willingness to trade off on treatment attributes such as benefits and risks.
- However, heterogeneity in preferences is common. Current approaches provide insights into the "who" and "what" of preference heterogeneity, but do not provide information on "why".
- Understanding the rationale behind preference heterogeneity is increasingly valuable as treatment options proliferate and the use of PP information by decision-making concurrently increases.

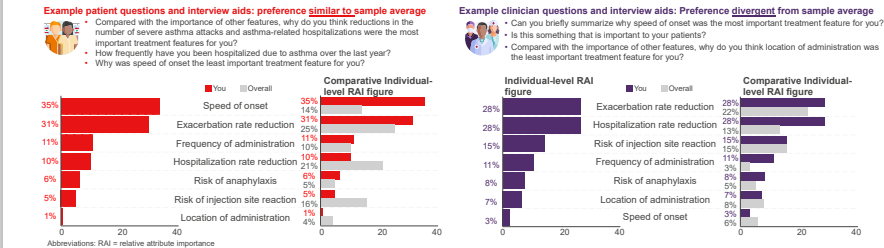
Objective

- We present a novel qualitative post-DCE interview method helpful in elucidating individuals' rationale for stated preferences and deriving rich insights into the drivers of the heterogeneity often observed in stated preference studies.

Methods

- A DCE was fielded among a sample of patients with severe uncontrolled asthma (n=300) and asthma-treating clinicians (n=247) in the US. Details on the design and results are reported elsewhere.¹
- Individual 1-hour online qualitative interviews were conducted with a subset of DCE respondents, after survey completion.
- Individual-level relative attribute importance (RAI) scores were generated for each DCE respondent using mixed logit estimates.
- Interview participants were selected if their individual-level RAI scores were either highly divergent from (target 75%) or characteristic of a typical response pattern (target 25%) compared with the RAI rank and preference directions of the overall sample.
- Interviews followed a semi structured guide. Participants were shown a figure of their individual-level RAI scores, then a figure comparing their individual-level scores with the median score for each attribute RAI. Participants were probed on the reasons for their preferences (Figure 1).

Figure 1. Example Participant Selection and Individual-level RAI Outputs



Results

Figure 2. Main Study Overall Sample RAI by Subgroup¹

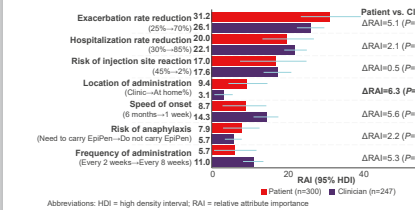
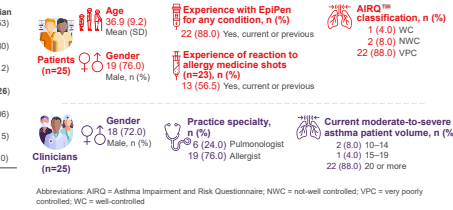


Figure 3. Descriptive Characteristics



Thermo Fisher
SCIENTIFIC

PCR170

Results (cont.)

Figure 4. Key Rationale for Asthma Treatment Preferences

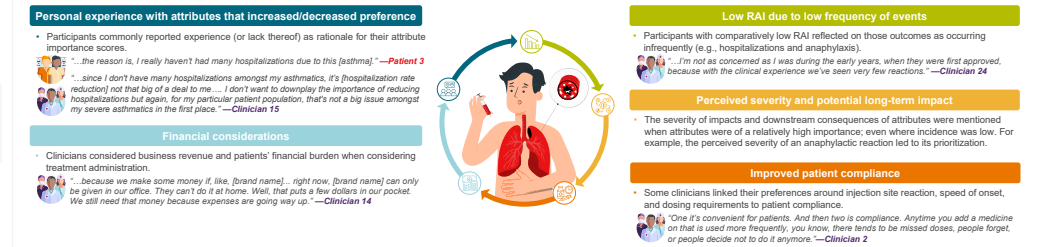
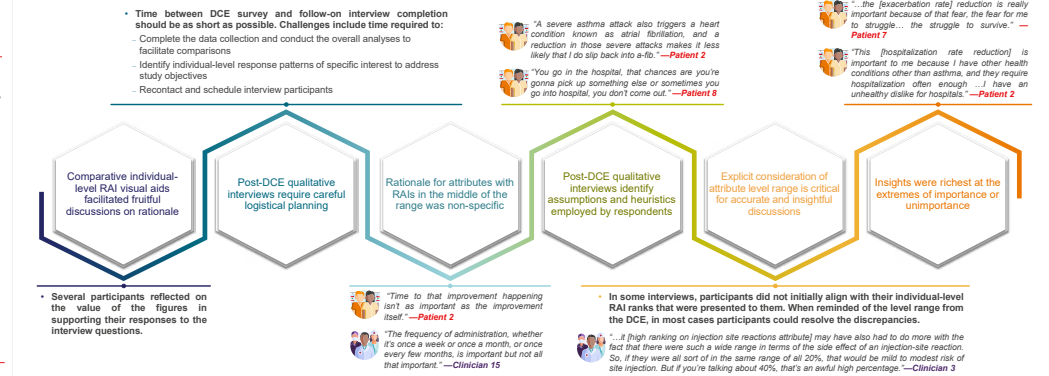


Figure 5. Post-DCE Qualitative Interviews: Methodological Learnings



Limitations

- Participants' opt-out behaviors were not analyzed or considered, future studies might probe the rationale behind opt-out behaviors.
- The post-DCE interviews took place several months after completion of the DCE survey, ideally the interviews would be conducted within a few days to weeks of completion of the DCE.

Conclusions

- Post-DCE qualitative interviews provide valuable insights into the rationale behind treatment preferences estimated using quantitative approaches.
- These data may be particularly valuable when preferences are highly heterogeneous and/or the rationale for specific preferences is poorly understood.
- The results may help explain heterogeneity and/or identify relevant covariates that may be evaluated in quantitative models exploring heterogeneity.

References

1. Ross et al. J Asthma. 2024;61(12):1746-58.

Disclosures and Acknowledgments

HG, HC, and MR are employees of PPD™ Evidera™ Patient-Centered Research, Thermo Fisher Scientific. Funding provided by Thermo Fisher Scientific. Editorial and graphic design support were provided by Karissa Calera and Kavitha Nakayira of Thermo Fisher Scientific.



ISPOR (NA) 2025

13–16 May 2025 | Montreal, Quebec, Canada

Poster Title

Mind the Gap: Delays between Vaccine Licensure and NITAG Recommendation in the US vs. Europe

Objective

- Given the importance of vaccine uptake for the prevention of communicable diseases, the goal of this project was to compare the delay between vaccine licensure and NITAG evaluation in the US vs. Europe.

AUTHORS

Samantha James¹
Ruth Chapman²
Kassandra Schaible³
Anna D'Ausilio⁴

AFFILIATION

¹Thermo Fisher Scientific, Paris, France
²Thermo Fisher Scientific, London, UK
³Thermo Fisher Scientific, Waltham, MA, USA
⁴Thermo Fisher Scientific, Milan, Italy

PRACTICE AREA

PPD™ Evidera™ Health
Economics & Market Access,
Thermo Fisher Scientific

CLIENT

INTERNAL



Thermo Fisher
SCIENTIFIC

HTA95

Mind the Gap: Delays between Vaccine Licensure and NITAG Recommendation in the US vs. Europe

Samantha James¹, Ruth Chapman², Kassandra Schaible³, Anna D'Ausilio⁴

¹Thermo Fisher Scientific, Paris, France; ²Thermo Fisher Scientific, London, UK; ³Thermo Fisher Scientific, Waltham, MA, USA; ⁴Thermo Fisher Scientific, Milan, Italy

Background

- The pathway to market access for vaccines differs from pharmaceuticals and varies significantly based on geography and the involvement and unique practices of National Immunization Technical Advisory Groups (NITAGs).¹
- In the US, the Advisory Committee on Immunization Practices (ACIP) is the NITAG in charge of evaluating and providing recommendations of vaccines.
- Most countries in Europe also have a NITAG, which tend to follow a formal evaluation pattern similar to that of health technology assessment bodies (HTAB), but the involvement of HTAB is not consistent across geographies,² and some countries—such as the US—do not have an HTAB.
- The non-standard approach from one country to another creates delays between vaccine licensure and recommendations, which further delays consumer access to these imperative public health technologies.

Objectives

Given the importance of vaccine uptake for the prevention of communicable diseases, the goal of this project was to compare the delay between vaccine licensure and NITAG evaluation in the US vs. Europe.

Methods

A targeted review of available literature was conducted to determine the time from vaccine licensure in the US to ACIP review and the subsequent publication of the decision by the Centers for Disease Control and Prevention (CDC).

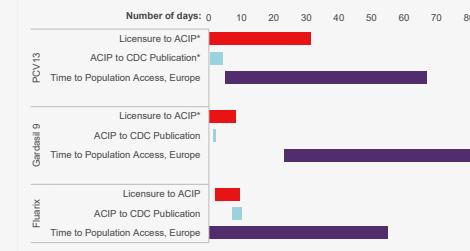
The search was conducted using PubMed®, the CDC's Morbidity and Mortality Weekly Report, and published ACIP recommendations.

The vaccines included in the review were limited to a pneumococcal vaccine (13-valent pneumococcal conjugate vaccine [PCV13]/Prennar 13), a human papillomavirus (HPV) vaccine (Gardasil 9), and a quadrivalent influenza vaccine (Fluarix Quadrivalent). These vaccines were selected to facilitate comparison with a previously published analysis on the delay between licensure and NITAG recommendation in Europe.¹

Results

- Nineteen publications were included that report time to ACIP recommendation for the vaccines of interest (Table 1).
- In the US, ACIP recommendations were typically made within 10 months of vaccine licensure, except for one instance of a delay of 2.6 years for recommendation of expanded use of PCV13 in older adults. In two cases, ACIP issued guidance prior to licensure, recommending expanded use for PCV13 and Gardasil 9 for high-risk populations a respective 4 years and 10 months ahead of regulatory approval.
- Recommendations made by ACIP were generally published by the CDC within 4 months for PCV13 and Gardasil 9, and within 10 months for Fluarix. Overall, the time from licensure to published NITAG decision did not exceed 2 years for any of the included vaccines.
- In contrast, a study published by Laigle and colleagues demonstrated that the time from licensure to NITAG recommendation in Europe exceeded 6 years in almost 50% of the 28 countries studied (including France, Germany, Italy, and the UK), with only four countries (Cyprus, Estonia, Lithuania, and Malta) reporting a delay of fewer than 2 years (Figure 1).¹ Most of the countries with a delay of 6 years or more involved parallel NITAG review with an HTAB.
- Of the countries with a delay of fewer than 2 years in Europe, all but Estonia lack an HTAB, similar to the US.

Figure 1. Time from Licensure to NITAG Recommendation in the US vs. Europe for Select Vaccines



*Recommendations received prior to licensure were omitted from the figure. Abbreviations: ACIP = Advisory Committee on Immunization Practices; CDC = Centers for Disease Control and Prevention; PCV13 = 13-valent pneumococcal conjugate vaccine.

Table 1. Time from Licensure to NITAG Recommendation in the US for Select Vaccines

FDA Licensed Population(s)	ACIP			CDC		
	Date	Date	Delay ^a	Date	Delay ^b	ACIP/CDC Population(s)
PCV13/Prennar 13 (Wyeth Pharmaceuticals Inc., a subsidiary of Pfizer Inc.)^c						
Children (6 weeks to 5 years of age) ^d	24 FEB 2010 ^d	24 FEB 2010 ^d	0 days	12 MAR 2010 ^d	16 days	• All children 2–59 months of age ^d • Children 60–71 months of age with specific comorbidities ^d
Children/adolescents (6–17 years of age) ^e	25 JAN 2013 ^e	20 FEB 2013 ^e	26 days	28 JUN 2013 ^e	4.3 months	• Children 60–18 years of age with specific risk factors ^e
Adults (18–49 years of age) ^f	12 JUL 2016 ^f	20 JUN 2012 ^g	-4 years ^f	12 OCT 2012 ^g	3.7 months	• Adults >19 years of age with specific risk factors ^h
Older adults (≥50 years of age) ^g	30 DEC 2011 ^g	13 AUG 2014 ^g	2.6 years	19 SEP 2014 ^g	37 days	• All adults ≥65 years of age ^h
Gardasil 9 (Merck Sharp & Dohme Corp.)						
Girls and women (9–26 years of age); boys (9–15 years of age) ^h	10 DEC 2014 ^h	FEB 2015 ⁱ	67 days	27 MAR 2015 ⁱ	40 days	• Females from 11 or 12–26 years of age ⁱ • Males from 11 or 12–21 years (26 years for males with specific risk factors) ^{i,j}
Boys and men (16–26 years of age) ^h	14 DEC 2015 ^h	FEB 2015 ⁱ	-10 months ^f	27 MAR 2015 ⁱ	40 days	• Males from 11 or 12–21 years (26 years for males with specific risk factors) ^{i,j}
Women and men (27–45 years of age) ^h	5 OCT 2018 ^h	JUN 2019 ^h	8.3 months	16 AUG 2019 ^h	62 days	• Selective use via shared clinical decision-making for some adults 27–45 years of age ^h
Fluarix Quadrivalent (GlaxoSmithKline Biologicals)						
Children and adults (age ≥3 years of age) ^h	14 DEC 2012 ^h	20 FEB 2013 ^h	68 days	20 SEP 2013 ^h	7 months	• Licensed population
Children (age 6–35 months of age) ^h	11 JAN 2018 ^h	25 OCT 2018 ^h	9.5 months	23 AUG 2019 ^h	10 months	• Licensed population

^a From the date of licensure by the US FDA
^b From the date of ACIP decision
^c Pneumococcal indication only
^d Chronic heart disease, chronic lung disease, diabetes mellitus, cerebrospinal fluid leak, cochlear implant, alcoholism, chronic liver disease, cirrhosis, cigarette smoking, sickle cell disease/other hemoglobinopathy, congenital or acquired asplenia, congenital or acquired immunodeficiency, HIV infection, chronic renal failure, nephrotic syndrome, leukemia, lymphoma, Hodgkin disease, generalized malignancy, iatrogenic immunosuppression, solid organ transplant, and multiple myeloma
^e Immunocompromising conditions, functional or anatomic asplenia, cerebrospinal fluid leaks, or cochlear implants
^f Recommendation made by ACIP prior to licensure
^g Recommendation has since been modified to adults ≥65 years of age with specific risk factors (routine use) or who are otherwise healthy (selective use via shared clinical decision-making)
^h Men who have sex with men and immunocompromising conditions (including HIV infection)
ⁱ Abbreviations: ACIP = Advisory Committee on Immunization Practices; CDC = Centers for Disease Control and Prevention; PCV13 = 13-valent pneumococcal conjugate vaccine

Conclusions

- The time between vaccine licensure and NITAG recommendation is substantially shorter in the US than in almost all European countries. This is likely due to parallel or subsequent reviews from an HTAB and NITAG in most European countries.
- Expanded use of vaccines typically resulted in a quicker NITAG recommendation, even prior to licensure in some cases.
- Information sharing across NITAGs and HTABs as well as standardized frameworks would be beneficial to avoid delays.
- Joint clinical appraisal evaluation of vaccines may help reduce the time between licensure and recommendation in European countries.
- Development of standard approaches with respect to vaccine assessment and recommendation may help lead to quicker global consumer access to vaccines.

References

- Laigle V, et al. *Vaccine*. 2021;39(36):5706-18.
- Largeron N, et al. *Value Health*. 2024;27(10):1318-27.
- US Food and Drug Administration. February 24, 2010 Approval Letter - Prennar 13 [6 weeks to 5 years of age]. 2010. <https://www.fda.gov/biologics/blood/blood/vaccines/vaccines/ApprovedProducts/UCM477341.pdf>.
- Centers for Disease Control and Prevention (CDC). *MMWR Morb Mortal Wkly Rep*. 2010;59(9):258-61.
- US Food and Drug Administration. January 25, 2013 Approval Letter - Prennar 13 [6 to 17 years of age]. 2013. <https://www.fda.gov/biologics/blood/blood/vaccines/vaccines/ApprovedProducts/UCM462278.pdf>.
- Maher E, et al. *MMWR Morb Mortal Wkly Rep*. 2019;68(52):668-702.
- US Food and Drug Administration. December 14, 2012 Approval Letter - Fluaria Quadrivalent [3 years of age]. 2012. <https://www.fda.gov/biologics/blood/blood/vaccines/vaccines/ApprovedProducts/UCM32484.htm>.
- Advisory Committee on Immunization Practices (ACIP). Advisory Committee on Immunization Practices (ACIP) summary report 1 - February 20-21, 2013. Atlanta, Georgia: 2013. <https://stacks.cdc.gov/view/cdc/75321>.
- US Food and Drug Administration. December 30, 2011 Approval Letter - Prennar 13 [60 months of age]. 2011. <https://www.fda.gov/biologics/blood/blood/vaccines/vaccines/ApprovedProducts/UCM285434.htm>.
- Tomczyk S, et al. *MMWR Morb Mortal Wkly Rep*. 2014;63(37):822-5.
- US Food and Drug Administration. December 10, 2014 Approval Letter - GARDASIL 9 [9 to 26 years of age; boys: 9 to 15 years of age]. 2014. <https://www.fda.gov/biologics/blood/blood/vaccines/vaccines/ApprovedProducts/UCM426520.htm>.
- Grusakov LA, et al. *MMWR Recomm Rep*. 2015;64(11):300-4.

Disclosures
S.J., R.C., K.S., and A.D. are employees of PPD™ Evidera™ Health Economics & Market Access, Thermo Fisher Scientific. Poster development was funded by Thermo Fisher Scientific.

Acknowledgments
Editorial and graphic design support were provided by Michael Orosi, Shan Berger, and Kaushar Nakayama of Thermo Fisher Scientific.



Poster Title

Exploring Treatment Preferences for Rare Diseases: A Systematic Review of Quantitative Preference Studies

Objective

- Patient preference information (PPI) is increasingly important throughout the medical product lifecycle. PPI can provide valuable insights into the relative importance that patients and their families place on benefits and risks associated with novel treatments, which can inform regulatory and clinical decision-making.
- Using standard quantitative preference elicitation methods to generate robust PPI can be challenging with small patient populations, particularly when there are few existing treatments and uncertainty around clinical evidence.
- To understand the state of practice, systematic review of the literature was conducted to identify studies using quantitative preference elicitation methods.

<p>AUTHORS</p> <p>Christine Michaels-Igbokwe¹</p> <p>Malavika Venkatraman²</p> <p>Keila Meginnis²</p> <p>Erica Visintin²</p>	<p>AFFILIATION</p> <p>Zaneta Balantac³</p> <p>Ilene Hollin⁴</p> <p>Norah Crossnohere⁵</p>	<p>PRACTICE AREA</p> <p>PPD™ Evidera™ Patient-Centered Research, Thermo Fisher Scientific</p>	<p>CLIENT</p> <p>INTERNAL</p> <p>Temple University, Pennsylvania, USA</p> <p>Ohio State University, USA</p>
---	---	--	--



PCR110

Exploring Treatment Preferences for Rare Diseases: A Systematic Review of Quantitative Preference Studies

Christine Michaels-Igbokwe¹, Malavika Venkatraman², Keila Meginnis², Erica Visintin², Zaneta Balantac³, Ilene Hollin⁴, Norah Crossnohere⁵
¹Thermo Fisher Scientific, St Laurent, Quebec, Canada; ²Thermo Fisher Scientific, London, UK; ³Thermo Fisher Scientific, Waltham, MA, USA; ⁴Temple University, Pennsylvania, USA; ⁵Ohio State University, Ohio, USA

Objectives

- Patient preference information (PPI) is increasingly important throughout the medical product lifecycle. PPI can provide valuable insights into the relative importance that patients and their families place on benefits and risks associated with novel treatments, which can inform regulatory and clinical decision-making.
- Using standard quantitative preference elicitation methods to generate robust PPI can be challenging with small patient populations, particularly when there are few existing treatments and uncertainty around clinical evidence.
- To understand the state of practice, systematic review of the literature was conducted to identify studies using quantitative preference elicitation methods to assess preferences for treatments for rare diseases.

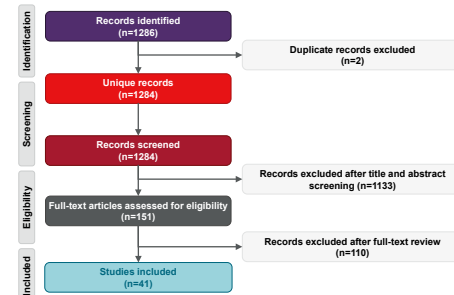
Methods

• A systematic search of EMBASE and MEDLINE was conducted in March 2023 to identify English-language articles reporting results of primary research. Studies were double-screened, and extraction was completed following a prespecified template.

Results

• A total of 1,286 citations were identified and screened based on title and abstract. Following full-text review of 151 studies, 41 met inclusion criteria and were included in the review (Figure 1).

Figure 1. PRISMA Diagram

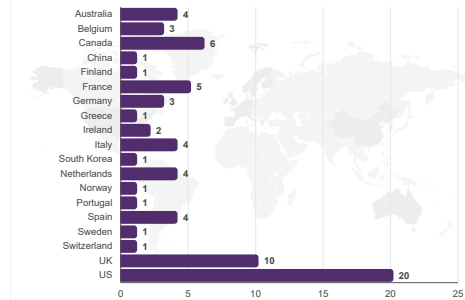


Abbreviations: PRISMA = preferred reporting items for systematic reviews and meta-analyses

Results (cont.)

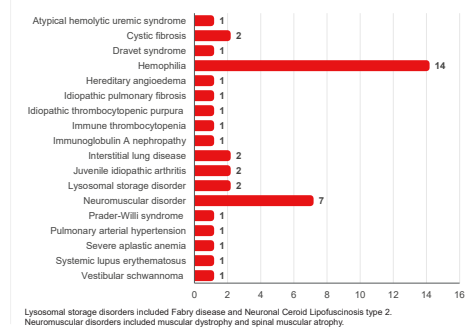
• Studies were published between 2005 and 2023, with 54% (n=22) published in 2020 or later. The majority of studies (n=32, 78%) were conducted in a single country; nine studies (22%) were conducted in two or more countries (Figure 2).

Figure 2. Study Countries



• A variety of health conditions were represented in the included preference studies (Figure 3). Overall, 30 studies (73%) assessed preferences for treatment of genetic disorders. Of these, treatment for hemophilia was the subject of 14 studies (47%). Seven studies (23%) assessed preferences for the treatment of neuromuscular disorders, such as muscular dystrophy and spinal muscular atrophy.

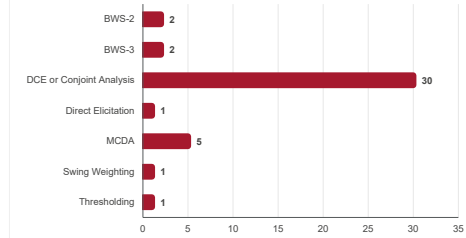
Figure 3. Number of Studies by Disease Area



Lysosomal storage disorders included Fabry disease and Neuronal Ceroid Lipofuscinosis type 2. Neuromuscular disorders included muscular dystrophy and spinal muscular atrophy.

• Discrete choice experiments (DCEs) were the most frequently used method of preference elicitation (n=31, 75%), followed by multicriteria decision analysis (MCDA; n=5, 12%) and best-worst scaling (BWS; n=4, 10%). Swing weighting, direct elicitation, and thresholding were used in one study each (Figure 4).

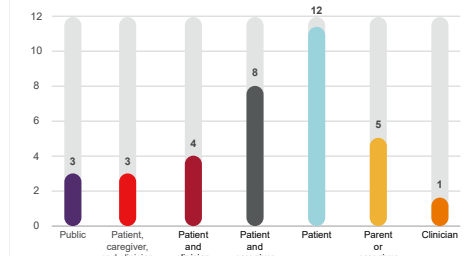
Figure 4. Preference Elicitation Methods*



Abbreviations: BWS = best-worst scaling, DCE = discrete choice experiment, MCDA = multicriteria decision analysis
 *More than one method used in one study

- Instrument design was informed by a review of published literature or clinical data in half of included studies (n=20, 49%).
- One quarter of studies engaged with community stakeholders (patients or caregivers) (n=5, 12%) or conducted qualitative research with patients (n=4, 10%) or parents (n=1, 2%).
- Respondent populations included patients (n=28, 68%), parents/caregivers (n=14, 34%), clinicians (n=10, 24%), and the general population (n=3, 7%) (Figure 5).

Figure 5. Respondent Populations in Choice-based Preference Studies*

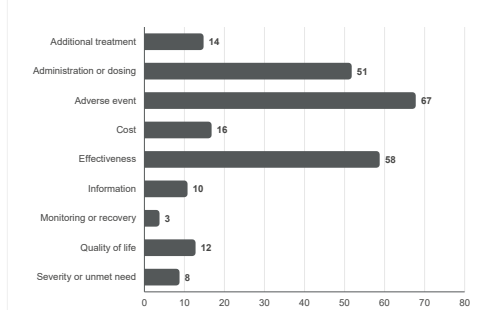


*Excludes studies using MCDA (multicriteria decision analysis)

• Across choice-based preference studies, studies with respondent populations that included patients had a mean of 164 patients (range: 9–1,542, median: 94). Studies that included parent/caregiver respondents a mean of 104 parents/caregivers (range: 19–468, median: 84).

• The number of attributes included in each preference study ranged from 1–13 (mean=6). Across all studies, 239 attributes were included. Of these, the most frequent types of attributes were adverse events (n=67), treatment efficacy or effectiveness (n=58), and items related to treatment administration or dosing (n=51) (Figure 6).

Figure 6. Typology of Included Attributes



Conclusions

- DCEs are the most frequently used method of elicitation when assessing preferences for treatment for rare diseases.
- Preference studies are often designed without patient or parent/caregiver input and therefore may not include attributes of treatment that are most meaningful to patients and families affected by rare diseases.
- Exploration and adoption of individual-level preference elicitation methodologies that are suitable for implementation with smaller sample sizes can support increasing adoption and utilization of PPI among rare disease populations.

Disclosures

CM, MV, KM, and EV are employees of PPD™ Evidera™ Patient-Centered Research, Thermo Fisher Scientific. ZB was an employee of Thermo Fisher Scientific at the time this study was conducted. Poster development was funded by Thermo Fisher Scientific.

Acknowledgments

Editorial and graphic design support were provided by Karissa Mariete Catara and Kawthar Nakayima of Thermo Fisher Scientific.



Poster Title

Stability of a Large Language Model for Data Extraction in Systematic Literature Reviews

Objective

- This study aimed to evaluate the reproducibility and reliability of LLM-extracted data when considering variations with the same user or different users from two geographic locations. Additionally, the accuracy of LLM-extracted data was compared with manual human extractions from a previously conducted traditional SLR.

Stability of a Large Language Model for Data Extraction in Systematic Literature Reviews

Aiswarya Shree¹, Mariana Farraia², Carolina Casañas i Comabella³, Allie Cichewicz⁴
¹Thermo Fisher Scientific, Bengaluru, India; ²Thermo Fisher Scientific, Ede, Netherlands; ³Thermo Fisher Scientific, London, UK; ⁴Thermo Fisher Scientific, Waltham, MA, USA



MSR105

Background

- Artificial Intelligence (AI) has been extensively explored in systematic literature reviews (SLRs) to save time and reduce human error, including the ability of large language models (LLMs) to perform data extraction has been tested.^{1,2}
- We previously reported the high accuracy (84%, range: 66% to 96%) of an LLM for data extraction in an SLR of randomized controlled trials (RCTs).³ However, we noted variations in responses when the same prompts were used on different days.
- Recently, the UK's National Institute for Health and Care Excellence (NICE) released a position statement on the use of AI for evidence generation, raising concerns about the reproducibility of AI, particularly regarding automated data extraction.⁴
- To our knowledge, there is limited evidence on the reproducibility and reliability of LLMs for data extraction, particularly given that some response variability is expected with these models and the impact on the trustworthiness of LLM-extracted data has not yet been characterized.

Objectives

- This study aimed to evaluate the reproducibility and reliability of LLM-extracted data when considering variations with the same user or different users from two geographic locations. Additionally, the accuracy of LLM-extracted data was compared with manual human extractions from a previously conducted traditional SLR.

Methods

- Three previously developed³ one-shot prompts were used to extract 29 variables from five RCTs on atopic dermatitis⁵⁻⁸ (Table 1).

Table 1. Variables for data extraction

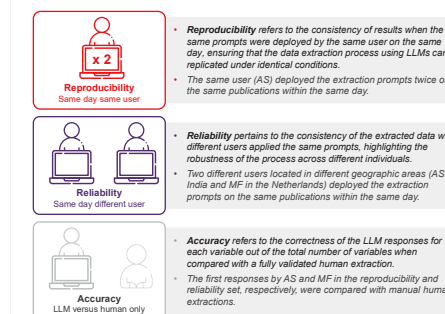
Category	Variable Type	Variables
Study characteristics	Free-text fields	Author, year, trial name, phase, population description, intervention, comparator, inclusion criteria, exclusion criteria, and overall sample size
	Numeric fields	Author, year, treatment arm
Patient characteristics	Free-text fields	Author, year, treatment arm, analysis population, time point
	Numeric fields	Sample size, mean age, male sex (%), comorbidities (%), disease severity (%)
Outcomes	Free-text fields	Author, year, treatment arm, analysis population, time point
	Numeric fields	Sample size, mean CFB in DLQI score, EASI 75 (%), POEM (%), treatment discontinuation (%), serious adverse events (%)

Numeric fields included binary, categorical, and continuous variables.
 Abbreviations: CFB = change from baseline; DLQI = Dermatology Life Quality Index; EASI = Eczema Area and Severity Index; POEM = Patient-Oriented Eczema Measure

- Two reviewers (AS and MF) used the same LLM prompts, with creativity set to 0, and the same publications to test data extraction by the LLM for reproducibility and reliability between responses (Figure 1) and accuracy compared with human extraction.
- The first set of LLM responses obtained by AS in the reproducibility test served as the reference and was compared with the second LLM extraction by AS (reproducibility) and MF (reliability). Reproducibility and reliability were calculated as the proportion of variables where LLM-extracted content was the same between responses, considering both content and formatting.
- Accuracy was calculated as the proportion of correctly extracted variables compared with the validated human extractions (conducted previously).³

Methods (cont.)

Figure 1. Methods to assess reproducibility, reliability, and accuracy



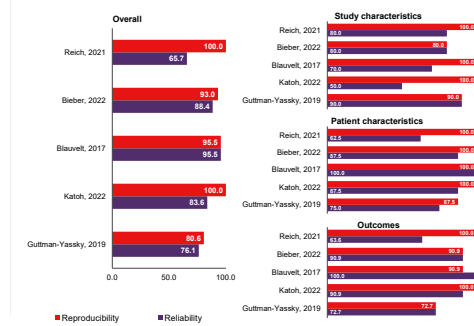
Abbreviations: AS = Aiswarya Shree; LLM = large language model; MF = Mariana Farraia

Results

Reproducibility

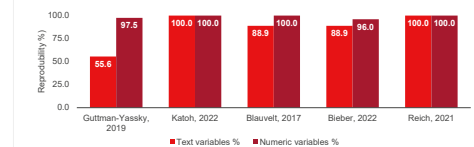
- Reproducibility of LLM responses to the data extraction prompts ranged from 80.6% to 100% (Figure 3).
- Reproducibility of responses for patient characteristics and outcome variables ranged from 88.5% to 100% and 72.7% to 100%, respectively.
- Reproducibility of responses for study characteristics varied from 80% to 100%.
- Reproducibility of responses for text vs. numeric variables are displayed in Figure 4.
- Reproducibility of responses for text variables was >85% in four studies, except for one study with 55.6%. Reproducibility for numeric variables was >95% in all studies.

Figure 3. Overall reproducibility and reliability; and by categories



Results (cont.)

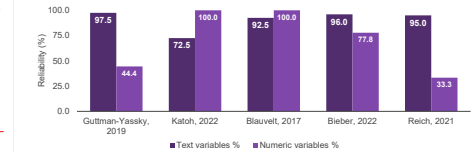
Figure 4. Reproducibility for text vs. numeric (binary, categorical, and continuous) variables



Reliability

- Reliability was lower than reproducibility, ranging from 65.7%–95.5% (Figure 3).
- Reliability for patient characteristics and outcome variables ranged from 62.5% to 100% and 63.7% to 100%, respectively.
- Reliability for study characteristics variables was the lowest, ranging from 50% to 90%.
- Reliability in text vs. numeric variables.
- None of the extractions were 100% reliable regarding text variables (range: 72.5% to 97.5%). Reliability of numeric variables varied considerably; responses for two studies had <50% reliability, while two others had 100% (Figure 5).

Figure 5. Reliability for text vs. numeric (binary, categorical, and continuous) variables



Accuracy

- Compared with validated human extractions, the LLM did not achieve an overall extraction accuracy of 100% for any publications in either test. The accuracy of LLM-extracted data was slightly higher with the reproducibility set of responses (79.1% to 98.5%) compared with the reliability set (74.6% to 97.7%).
- Accuracy with text variables: Overall accuracy remained consistent between response sets as the LLM captured the same underlying information for text variables. Despite slight variations in syntax, style, or length of response, this did not impact the accuracy, as the content of the LLM-extracted data aligned with the manual human reference extraction.
- Accuracy with numeric variables: Accuracy was adversely affected by the extraction of numeric fields, where discrepancies were observed between LLM responses and the human reference extraction.

Discussion

- This study highlights both the benefits and potential challenges of using LLMs for data extraction from RCTs. Our findings indicate high reproducibility rates, ranging from 80.6% to 100%, suggesting that LLMs can consistently replicate extraction of data under the same user conditions. However, reliability, which evaluates consistency between different users, was lower despite using identical prompts (range: 65.7% to 95.5%).

Discussion (cont.)

- Overall, the accuracy of LLMs for data extraction was high.
- The observed variations in text responses did not negatively affect the overall accuracy of data extraction.
- However, in addition to formatting inconsistencies, the extraction of numeric fields created discrepancies that led to errors that affected accuracy.
- SLRs require meticulous data extraction processes, often involving multiple extractors, followed by thorough data validation. This process may take several weeks (depending on study volume) and is susceptible to human error (up to 50%).^{10,11} Using an LLM-based extraction approach may result in faster, more consistent and reliable results, particularly over time and across different users, and potentially reduce human error and increasing extraction quality.
- Additional research is needed to understand and mitigate the factors contributing to variation in LLM-extracted data, including developing techniques to improve the consistency of numeric data extraction and to reduce the impact of stochastic elements in LLMs.

Limitations

- Our findings may not be generalizable to other types of studies beyond RCTs, as different study designs and reporting of outcomes may introduce additional challenges.
- This study tested a small sample of publications. Larger volumes of data might exhibit higher variations in LLM responses, particularly if variables are reported more heterogeneously across publications.
- One-shot prompting was used; a more iterative prompting approach might achieve better accuracy, but could also introduce further variations between users, potentially impacting reproducibility and reliability.
- Testing was conducted at a specific time point (November 2024). Given how quickly LLMs are evolving, future testing may demonstrate different or improved results.
- Only one LLM was tested. Other LLMs may yield different results, and further research is needed to compare the performance of various models in data extraction tasks.

Conclusions

- When using LLMs for data extraction in SLRs, reproducibility was generally high, but reliability was affected by user interaction, with variations between different users leading to discrepancies, particularly for numeric data. These findings highlight the need for human validation of LLM-extracted data to ensure data quality.
- While variations in text responses are expected with LLMs and the impact on overall accuracy was minimal, numerical discrepancies highlight the need for human oversight such that AI-driven extractions must still undergo human validation to ensure data accuracy.
- Transparent reporting of AI-assisted methods in SLRs is crucial to contextualize results and maintain scientific rigor. Our findings underscore the importance of addressing the concerns raised by NICE's AI position statement and are important for informing future Cochrane guidance on the applications of AI in SLRs.

References

- Koneti et al. *Res Synth Methods*. 2024;15(5):818-824. 2. Gantelmer et al. *Res Synth Methods*. 2024;15(4):578-589. 3. Shree et al. *Evidence Science Inc.*. 2024; 4. NICE. <https://www.nice.org.uk/about/what-we-do/our-research/working-with-ai-in-evidence-generation--nice-position-statement> 2025. 5. Bieber et al. *Br J Dermatol*. 2022;187(3):338-352. 6. Blauvelt et al. *Lancet*. 2017;389(10085):2291-2303. 7. Reich et al. *Lancet*. 2021;397(10260):2169-2181. 8. Gutman et al. *J Am Acad Dermatol*. 2019;80(4):613-621. e6. 9. Katoh et al. *Dermatol Ther*. 2023; 13(1):221-234. 10. Tang et al. *Postgrad Med J*. 2025;sgae195. 11. Mathes et al. *BMC Med Res Methodol*. 2017;17:1-8.

Disclosures

AS, MF, CDC, and AG are employees of PPD™ Evidera™ Health Economics & Market Access, Thermo Fisher Scientific.

Acknowledgments

Editorial and graphic design support were provided by Sean Smith and Kaveri Nayakuma of Thermo Fisher Scientific.

AUTHORS

Aiswarya Shree¹ Allie Cichewicz⁴
 Mariana Farraia²
 Carolina Casañas
 Comabella³

AFFILIATION

¹Thermo Fisher Scientific, Bengaluru, India
²Thermo Fisher Scientific, Ede, Netherlands
³Thermo Fisher Scientific, London, UK
⁴Thermo Fisher Scientific, Waltham, MA, USA

PRACTICE AREA

PPD™ Evidera™ Health Economics & Market Access, Thermo Fisher Scientific

CLIENT

INTERNAL



Poster Title

Leveraging Real-world Data vs. External Trials to Inform Comparative Efficacy within Single-arm Trial-based NICE Submissions

Objective

- This study reviewed the acceptance of RWD vs. external trials to inform comparative efficacy in SAT-based HTA submissions.

AUTHORS

Sherry Wu
Eileen Zhang
Denise Zou

AFFILIATION

Thermo Fisher Scientific, Waltham, MA, USA

PRACTICE AREA

PPD™ Evidera™ Health Economics & Market Access, Thermo Fisher Scientific

CLIENT

INTERNAL



Leveraging Real-world Data vs. External Trials to Inform Comparative Efficacy within Single-arm Trial-based NICE Submissions

Sherry Wu¹; Eileen Zhang¹; Denise Zou¹
¹Thermo Fisher Scientific, Waltham, MA, USA

Background

- Single-arm trials (SATs) have increasingly been used to support oncology appraisals by health technology assessment (HTA) bodies, driven by the ethical and practical challenges of conducting randomized controlled trials (RCTs) involving patients with specialized treatment needs.
- In the absence of direct comparative data in SATs, evidence may be obtained from external clinical trials and/or real-world data (RWD) to inform indirect treatment comparisons (ITCs).
- In consideration of the available external comparator data, statistical approaches including matching-adjusted indirect comparisons (MAIC), simulated treatment comparison (STC), or propensity score matching (PSM) may be employed.
- Currently, there are no explicit HTA guidelines for generating comparative evidence for SAT-based submissions.

Objectives

This study reviewed the acceptance of RWD vs. external trials to inform comparative efficacy in SAT-based HTA submissions.

Methods

- Oncology SAT-based appraisals from the National Institute for Health and Care Excellence (NICE) between May 2017 and May 2022 were reviewed.
- Full-text screening of committee papers and technology appraisal guidance was conducted by a single investigator, and the extracted data were validated by a second investigator.
- The review focused on identification of the ITC approach used to derive comparative efficacy and the related committee commentaries.

Results

- Of the 31 submissions reviewed, 58% (18/31) used external trials only to derive comparative efficacy, 29% (9/31) used RWD only, and 13% (4/31) used both, with the variance driven by external data availability, limitations, and relevance (Figure 1). Half of these sources were deemed fit-for-purpose by the committee.

Results (cont.)

Figure 1. Type of Evidence Used in NICE Submissions

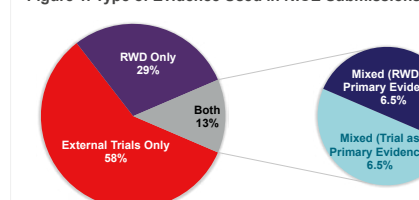
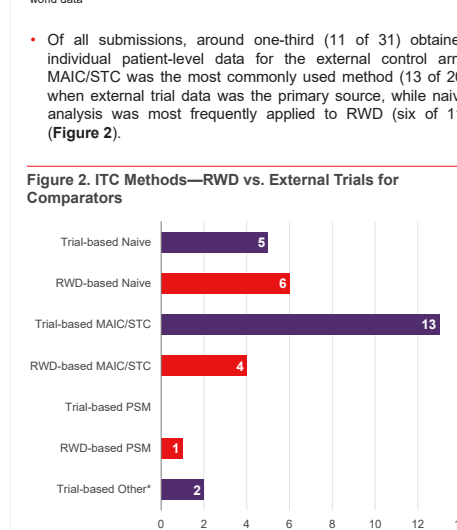


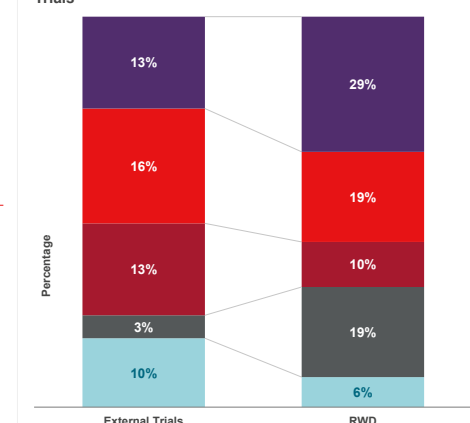
Figure 2. ITC Methods—RWD vs. External Trials for Comparators



Abbreviation: MAIC = matching-adjusted indirect comparison; PSM = propensity-score matching; RWD = real-world data; STC = simulated treatment comparison
*Note: Both submissions with other ITC approaches used single-arm trials as the primary source, with one constructing the comparator arm from its own trial and the other submission from landmark analysis.

- Among the 13 submissions leveraging RWD, more than 60% (n=8) were accepted as valid evidence,¹⁻⁸ benefiting from the ability to mitigate uncertainty due to lack of direct comparison by allowing precise matching and covariate adjustment.
- Overall, criticisms of SAT-based submissions were primarily due to data limitations (42%) and insufficient comparability (35%). Other concerns included limited generalizability of the results (23%), inadequate covariate adjustments (23%), and inappropriate statistical methods (16%) (Figure 3).

Figure 3. Criticism on Quality of Evidence: RWD vs. External Trials



Abbreviations: ITC = indirect treatment comparison; RWD = real-world data

- In particular, the committee emphasized that unanchored MAIC does not effectively reduce uncertainty or bias, often favoring supplementation with naive analyses as benchmarks (52%, 16/31).
- Submissions using external trials as primary evidence were mainly criticized for lack of comparability with trial population and insufficient adjustments of population difference, followed by concerns about result generalizability to the market-authorized population.
- In contrast, RWD-based submissions were criticized for high uncertainty due to data limitations (e.g., small sample size, immaturity, short follow-up), followed by insufficient covariate adjustments in ITC analyses and lack of comparability to trial populations.

Conclusions

RWD has been increasingly used as an alternative to suboptimal trials for external control in SAT-based HTA submissions, offering more granularity and flexibility. HTA consensus on its appropriateness for external control remains low, with key discussions on whether the data are fit-for-purpose and the adequacy of covariate adjustments.

References

- National Institute for Health and Care Excellence. Ibrutinib for treating Waldenström's macroglobulinaemia (TA691). 2017. Accessed April 18, 2025. <https://www.nice.org.uk/guidance/ta691>
- National Institute for Health and Care Excellence. Brentuximab vedotin for treating relapsed or refractory systemic anaplastic large cell lymphoma (TA478). 2017. Accessed April 18, 2025. <https://www.nice.org.uk/guidance/ta478>
- National Institute for Health and Care Excellence. Nivolumab for treating relapsed or refractory classical Hodgkin lymphoma (TA462). 2017. Accessed April 18, 2025. <https://www.nice.org.uk/guidance/ta462>
- National Institute for Health and Care Excellence. Tisagenlecleumab for treating relapsed or refractory B-cell acute lymphoblastic leukaemia in people aged up to 25 years (TA554). 2018. Accessed April 18, 2025. <https://www.nice.org.uk/guidance/ta554>
- National Institute for Health and Care Excellence. Pembrolizumab for treating relapsed or refractory classical Hodgkin lymphoma (TA545). 2018. Accessed April 18, 2025. <https://www.nice.org.uk/guidance/ta545>
- National Institute for Health and Care Excellence. Brintanovab for treating acute lymphoblastic leukaemia in remission with minimal residual disease activity (TA589). 2019. Accessed April 18, 2025. <https://www.nice.org.uk/guidance/ta589>
- National Institute for Health and Care Excellence. Avelumab for untreated metastatic Merkel cell carcinoma (TA691). 2021. Accessed April 18, 2025. <https://www.nice.org.uk/guidance/ta691>
- National Institute for Health and Care Excellence. Brentuximab vedotin for treating relapsed or refractory mantle cell lymphoma (TA677). 2021. Accessed April 18, 2025. <https://www.nice.org.uk/guidance/ta677>

Disclosures

SW, EZ, and DZ are employees of PPD™ Evidera™ Health Economics & Market Access, Thermo Fisher Scientific. Poster development was funded by Thermo Fisher Scientific.

Acknowledgments

Editorial and graphic design support were provided by Michael Gross and Kawthar Nakayima of Thermo Fisher Scientific.



Biostatistics & Health Analytics

7–9 July 2025 | Aegina, Greece

Posters

Extrapolating Overall Survival Data: A Practical Guide to
Methods and Model Selection for Long-term Predictions



Biostatistics & Health Analytics 2025

7–9 July 2025 | Aegina, Greece

Poster Title

Extrapolating Overall Survival Data: A Practical Guide to Methods and Model Selection for Long-term Predictions

Objective

- This poster offers a summary of prevalent techniques for extrapolating OS, describes their foundational assumptions, and emphasizes practical aspects for choosing and validating these models.

AUTHORS

Michail Galanakis, MSc¹
Venediktos Kapetanakis, PhD²

AFFILIATION

¹Thermo Fisher Scientific, Chania, Greece
²Thermo Fisher Scientific, London, UK

PRACTICE AREA

PPD™ Evidera™ Health
Economics & Market Access,
Thermo Fisher Scientific

CLIENT

INTERNAL

Extrapolating Overall Survival Data: A Practical Guide to Methods and Model Selection for Long-term Predictions

¹Michail Galanakis, MSc; ²Venediktos Kapetanakis, PhD
¹Thermo Fisher Scientific, Chania, Greece; ²Thermo Fisher Scientific, London, UK

Background

- Overall survival (OS) serves as a vital endpoint in clinical studies, especially in oncology, as it represents the most important outcome for the patient. However, clinical trials frequently have restricted follow-up periods due to high costs, resulting in only partial survival data.
- Survival extrapolation is a critical component in the evaluation of long-term outcomes in biomedical research and health technology assessments.
- As decision-makers seek to understand the cost-effectiveness and potential benefits of new healthcare interventions over a lifetime horizon, the need for robust and reliable extrapolation methods is paramount.

Objectives

- This poster offers a summary of prevalent techniques for extrapolating OS, describes their foundational assumptions, and emphasizes practical aspects for choosing and validating these models.

Methods

- Choosing an extrapolation method depends on the data characteristics, trial context, and assumptions about the underlying survival process. Table 1 (see QR code) presents a short description of each method, when it is appropriate, key underlying assumptions, along with strengths and limitations.
- All methods are described for a single outcome and population (e.g., a single treatment arm) but can easily be extended by including covariates in the fitted models following the guidance and recommendations of the National Institute for Health and Care Excellence Decision Support Unit.¹

Model Selection and Validation

- Assessing model fit and selecting the most appropriate model is important and involves a staggered approach:
 - Goodness-of-fit within the observed data:**
 - Visual inspection: Overlay predicted vs. observed (Kaplan-Meier) curves—alignment over the entire follow-up period indicates good fit.
 - Statistical fit: Akaike Information Criterion, Bayesian Information Criterion—lower values indicate better fit.
 - Clinical plausibility of extrapolation:**
 - Expert opinion is required to validate long-term extrapolations based on clinical plausibility. Clinical experts should also advise on the existence of a cure fraction, the choice of a landmark point, and the effect of subsequent treatments (which can inform the knots for a piecewise modeling approach).
 - Comparison with external data: Whenever possible, compare long-term projections with historical or real-world data from similar patient populations.
 - Sensitivity analysis:**
 - Evaluate the influence of various credible models on long-term extrapolations.
 - Present a spectrum of plausible results instead of results from a single modeling approach.

Practical Recommendations

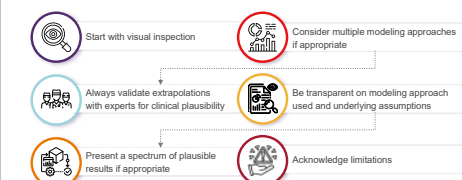
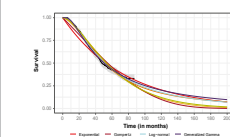


Figure 1. Standard Parametric Survival Models



Standard Parametric Survival Models (Figure 1): Time-to-event outcomes analyzed, assuming that the risk of the event follows a parametric distribution.

Bayesian Parametric Survival Analysis using Informative Priors (Figure 2): Integrating prior knowledge into parametric survival models using Bayesian methods, improving estimates and predictions.²

Presented at the 2nd Conference in Biostatistics & Health Analytics | 7-9 July 2025 | Aegina, Greece

Figure 3. Flexible Parametric Survival Models (Splines)

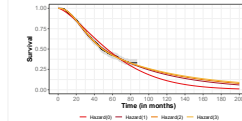


Figure 4. Mixture Cure Models

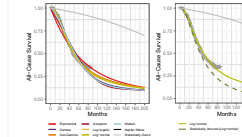
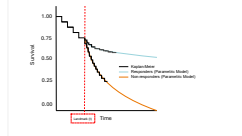


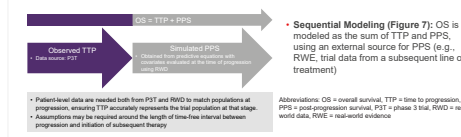
Figure 5. Landmark Analysis



Landmark Analysis (Figure 5): Involves selecting a specific point in time after the start of follow-up (the "landmark point"). All patients who have experienced the event of interest or have been censored before this landmark time are excluded from the analysis. For the remaining patients (those still at risk at the landmark time), survival analysis is then analyzed by resetting the clock from at the landmark point.

Piecewise Models (Figure 6): Instead of assuming a single underlying parametric distribution for the entire observation period, piecewise models define specific "change points" along the time axis. The survival function is then modeled separately within each of the intervals defined by these change points.

Figure 7. Sequential Modeling of OS as TTP + PPS



Conclusions

- Extrapolating OS data is a complex but essential objective in biostatistics, especially for health technology assessments.
- Each method for survival extrapolation has a unique role, and the choice of method should be guided by the specific context of the study, the maturity of the data, and the characteristics of the patient population.
- Understanding and reporting the strengths and limitations of these methodologies is essential for making informed decisions in health economic evaluations and biomedical research.
- Regardless of the choice of the modeling approach, clinical opinion is required to validate the results.

References

- NICE. NICE DSU Technical Support Document 14: Survival Analysis For Economic Evaluations. *Health Economics, Law and Ethics*. 2013. Available from: <https://www.nice.org.uk/guidance/TS14/resources/TS14-main-text>
- Sokal F, et al. *Value Health*. 2019;22(5):1012-1017.
- Lalor M, et al. *Pharmacoeconomics*. 2024;42:1073-1090.
- Rubner M, et al. *NICE DSU Technical Support Document 21: Flexible Methods for Survival Analysis*. 2022. Available from: <https://www.nice.org.uk/guidance/TS21/resources/TS21-main-text>

Disclosures

All authors are employees of PPD™ Evidera™ Health Economics & Market Access, Thermo Fisher Scientific. No honoraria were received by any author for their contribution to this work.

Acknowledgements: Editorial and graphic design support were provided by Michael Geiser and Heather Nakayama of Thermo Fisher Scientific.

To view Table 1, scan the QR code.

Extrapolating Overall Survival Data: A Practical Guide to Methods and Model Selection for Long-term Predictions

¹Michail Galanakis, MSc; ²Venediktos Kapetanakis, PhD
¹Thermo Fisher Scientific, Chania, Greece; ²Thermo Fisher Scientific, London, UK

Table 1. Modeling Approaches for Extrapolating OS along with Underlying Assumptions, Strengths and Limitations

Method	When Appropriate	Key Assumptions	Strengths	Limitations
Standard Parametric Survival Models Time-to-event outcomes analyzed, assuming that the risk of the event follows a parametric distribution	• Default approach. The seven parametric distributions provide a variety of shapes (monotonically increasing/decreasing/increasing and then decreasing), which often capture the underlying hazard function providing reliable extrapolations	• Hazard function follows a specific distribution (e.g., Exponential, Weibull, Gompertz, Log-logistic, Log-normal, and Generalized Gamma). Each distribution implies a particular shape for the hazard function, which is critical for long-term predictions	• Provide reliable extrapolations when the underlying hazard function is valid over time	• Extrapolations associated with significant uncertainty when trial data are immature • Cannot capture two or more turning points on the hazard function, in which case other modeling approaches may be preferable
Flexible Parametric Survival Models (Splines) Greater flexibility in capturing the shape of survival data while still allowing for the specification of underlying parametric forms (e.g., proportional hazards, proportional odds, probit model for event times based on a normal distribution)	• When standard parametric survival models fail to adequately capture the shape of the hazard curve • May be preferable to MCMs if the existence of "cure" is uncertain	• The chosen splines and placement of knots adequately capture the underlying hazard function • The log-hazard (or other transformed scales) will behave linearly beyond the final knot	• Improved fit to the observed data compared with standard parametric survival models	• Arbitrary choices for knots (number and location) that affect the reliability of the extrapolations • The "steep tail" may not provide sensible long-term extrapolations • Extrapolations, especially from PFS/OS with many knots, may be unreliable because they are based on a small fraction of the data after the last knot with only few events • Risk of overfitting
Mixture Cure Models The overall population is composed of two unobserved subgroups: a proportion of patients who are cured of their disease and follow general population mortality rates, and a group of uncured patients, whose excess mortality attributed to the disease is modeled using parametric distributions	• Plausibility of cure (e.g., observed plateau on KM with heavy censoring, and/or clinical expert opinion)	• Clinical justification for the expectation of a cure fraction • A fraction of the population is cured and experiences general population mortality (i.e., no excess mortality) • The excess risk of the uncured group follows a parametric distribution	• Explicitly models the existence of a cure fraction • Provides a more accurate representation of long-term survival in populations when a cure fraction is plausible • Can handle complex survival patterns by modeling cured and uncured populations separately • Can be used to estimate OS by assuming the same cure fraction as PFS	• Strong assumptions regarding cure, requiring biological evidence • Interpretation of cure: "Cured" patients are "cured" at baseline (backward interpretation). Non-MCMs can provide an alternative • Identifying the "cure" proportion can be challenging with limited follow-up
Bayesian Parametric Survival Analysis using Informative Priors Integrate prior knowledge into parametric survival models using Bayesian methods, improving estimates and predictions ²	• When relevant external data are available to form informative priors	• The prior information is relevant and accurately reflects the underlying survival process • The shape parameter is exchangeable between the studies	• Estimates the shape of the hazard function by incorporating external information in addition to the observed data from the index study. This is expected to improve survival extrapolations and reduce uncertainty because with an informative prior for the shape, the trial data can be used to better estimate the scale of the parametric distribution	• Requires robust external evidence for the informative prior (e.g., expert opinion, or the availability of a study with mature data on a similar population receiving the same/similar [e.g., in terms of mechanism of action] treatment, and the same outcome definition) • Assumption that the shape parameter of the outcome of interest is exchangeable between studies
Landmark Analysis Involves selecting a specific point in time after the start of follow-up (the "landmark point"). All patients who have experienced the event of interest or have been censored before the landmark time are excluded from the analysis. For the remaining patients (those still at risk at the landmark time), survival analysis is then analyzed by resetting the clock from at the landmark point	• Presence of a clinically meaningful landmark point (e.g., response status) • Early response predicts long-term outcome	• The chosen fixed landmark timepoint is clinically meaningful • Validity of patient status at the fixed landmark point • Conditional on the patients' status at the landmark point, survival after the landmark is independent of events that occurred before the landmark	• Addresses immortal-time bias • Leverages the fact that the survival data for patients who do not achieve response (e.g., non-responders) may have more mature data, which can be modeled robustly • Can be used to estimate survival for responders (which may have immature data) by applying an HR to the estimated curve for non-responders (typically with more mature data). The HR can be obtained from the literature	• Requires careful consideration on the choice of landmark point • Validity of patient status at the fixed landmark point is crucial as it can influence the results of the analysis • Requires sufficient sample size at landmark
Piecewise Models Instead of assuming a single underlying parametric distribution for the entire observation period, piecewise models define specific "change points" along the time axis. The survival function is then modeled separately within each of the intervals defined by these change points	• When standard parametric survival models fail to adequately capture the shape of the hazard curve • Clinical phases of disease/treatment (e.g., subsequent treatment may change the shape of curve)	• The selection of change point(s) is clinically meaningful • Validity of parametric survival distributions within each interval • Sufficient events/follow-up within each interval	• Increased flexibility to model complex survival patterns • Improved fit within each interval	• Arbitrary choices for change points (number and location) that affect the reliability of the extrapolations • Requires sufficient sample size and events within each interval • Complex interpretation • Extrapolations are based only on the last interval
Sequential Modeling OS is modeled as the sum of TTP and PPS, using an external source for PPS (e.g., RWE, trial data from a subsequent line of treatment)	• Immature data for OS but mature data for TTP • Available data sources to inform PPS • Clinical phases of disease/treatment	• PPS can be accurately modeled using RWE or trial data from a subsequent line of treatment • The sum of TTP and PPS provides a realistic estimate of OS	• Allows for more accurate extrapolation of OS by leveraging mature data for PPS from external sources • Can provide more insights into different phases of disease progression	• Requires reliable and relevant patient-level data for PPS to facilitate population matching at progression • Potential bias in RWE data • Complex to implement and validate

References

- NICE. NICE DSU Technical Support Document 14: Survival Analysis For Economic Evaluations. *Health Economics, Law and Ethics*. 2013. Available from: <https://www.nice.org.uk/guidance/TS14/resources/TS14-main-text>
- Sokal F, et al. *Value Health*. 2019;22(5):1012-1017.
- Lalor M, et al. *Pharmacoeconomics*. 2024;42:1073-1090.
- Rubner M, et al. *NICE DSU Technical Support Document 21: Flexible Methods for Survival Analysis*. 2022. Available from: <https://www.nice.org.uk/guidance/TS21/resources/TS21-main-text>

Disclosures

All authors are employees of PPD™ Evidera™ Health Economics & Market Access, Thermo Fisher Scientific.

Acknowledgements: Editorial and graphic design support were provided by Michael Geiser and Heather Nakayama of Thermo Fisher Scientific.

To view Table 1, scan the QR code.

International Society for Pharmacoepidemiology (ISPE)

22–26 August 2025 | Washington, DC, USA

Posters

Interpretable Machine Learning Models: Describing Treatment Patterns of Healthcare Providers Selecting Rheumatoid Arthritis Therapies to Evaluate Treatment Policies

Post-Approval Real-World Studies of RNA Therapies in the EU: What Can We Learn From Study Characteristics?

Beyond Comparison: Target Trial Emulation (TTE) Principles Strengthen Observational Single-arm (SA) Effectiveness Studies

Longitudinal, Regulatory-grade Registry Data Can Be Used to Emulate a Target Trial: An Application on Evaluating Treatments for Plaque Psoriasis

Policy Impact on Pregnancy Registry Participation and Reporting of Maternal/Infant Outcomes

Nuances of Risk Window Selection for Evaluation of Acute Adverse Events

Evolving Landscape of Post-marketing Surveillance (PMS) Studies in Japan Utilising Secondary Real-world Data (RWD)

Mandated Long-Term Follow-Up Studies of Patients Treated With Gene and Cell Therapies Approved in the European Union (EU): An Analysis of Commonalities and Variations

Review of Ongoing PASS Submitted to the HMA-EMA Catalogue of RWD Studies Evaluating Cancer as an Outcome

Vaccine-related Studies Listed in the Heads of Medicines Agencies-European Medicines Agency (HMA-EMA) Catalogue of Real-world Data Studies: What Can We Learn from Study Characteristics?

Methodological Considerations about Policy Impact on Studies of Pregnancy Registries and Maternal/Infant Outcomes

An Alternative to Non-responder Imputation for Study Participants with Missing Outcomes in Single-arm Drug-effectiveness Studies: A Case Study in the PPD CorEvitas Psoriasis Registry

Landscape and Study Design Characteristics of Prospective Observational Cohort Studies on Pregnant Women in China




Poster Title

Interpretable Machine Learning Models: Describing Treatment Patterns of Healthcare Providers Selecting Rheumatoid Arthritis Therapies to Evaluate Treatment Policies

Objective

- Demonstrate the strengths of interpretable ML models compared to black-box models for predicting treatment selection in rheumatoid arthritis (RA).

<p>AUTHORS</p> <p>Heather J. Litman¹ Zachary R Margolin¹</p> <p>Anton Matsson² Leslie R Harrold¹</p> <p>Yaochen Rao² Fredrik D Johansson²</p> <p>Jinming Wei²</p>	<p>AFFILIATION</p> <p>¹Thermo Fisher Scientific, Waltham, MA, USA</p> <p>²Chalmers University and University of Gothenburg, Gothenburg, Sweden</p>	<p>PRACTICE AREA</p> <p>PPD™ CorEvitas™ Clinical Registries, Thermo Fisher Scientific</p>	<p>CLIENT</p> <p>INTERNAL</p> <p>Chalmers University and University of Gothenburg</p>
--	---	--	--



A-175

Interpretable Machine Learning Models: Describing Treatment Patterns of Healthcare Providers Selecting Rheumatoid Arthritis Therapies to Evaluate Treatment Policies

Heather J. Litman¹, Anton Matsson², Yaochen Rao², Jinming Wei², Zachary R Margolin¹, Leslie R Harrold¹, Fredrik D Johansson²
¹Thermo Fisher Scientific, Waltham, MA, USA; ²Chalmers University and University of Gothenburg, Gothenburg, Sweden

Background

- With developments in computing power and artificial intelligence, there is more widespread implementation of complex machine learning (ML) black-box models, including random forest (RF) and extreme gradient boosting (XGB).
- However, explaining the modeling process is critical, especially in clinical outcome-related domains, such as evaluating treatment policies in healthcare.
- Rule-based interpretable ML models can lead to more transparency, facilitating comprehensive and understandability.

Objective

- Demonstrate the strengths of interpretable ML models compared to black-box models for predicting treatment selection in rheumatoid arthritis (RA)

Methods

PPD™ CorEvitas™ RA Registry

- The CorEvitas RA Registry is a US-based prospective, multicenter, observational, disease-based registry.
- RA patients typically initiate a conventional, synthetic disease-modifying anti-rheumatic drug (csDMARD); they initiate a biologic-targeted synthetic DMARD (btsDMARD) if the initial therapy fails.
- Between 2012 and 2021, 42,068 patients enrolled in the Registry, and 6,037 met the selection criteria, as follows:
 - Not having history of any DMARD use at registry enrollment
 - Initiating at least one btsDMARD while in the registry
 - Providing complete history of treatment use
- Disease activity was measured using the Clinical Disease Activity Index (CDAI), which consists of physician-reported joint count, global assessment, and patient's global assessment.
- The aim was to model the choice of DMARD therapy for patients with RA.^{1,2}

ML models

- Interpretable ML models are designed to make the reasoning process understandable to humans through transparency using two main types of models.⁴
- Rule-based models use series of "if-then" statements to create "rules."⁴
- Decision sets
 - Rule ensembles (RE) combine decision tree (DT) and modeling.⁵
 - Boosted rule sets (BRS) fit rules sequentially and aggregate with voting.⁶
- Decision lists
 - Greedy rule lists (RL) create rules by selecting the most significant feature at each step.
- DTs
 - Hierarchical shrinkage wrapper/rule trees (RTs) sequentially fit rules.⁷
- Linear models predict outcomes based on linear combinations of covariates.
- Logistic regression (LR) coefficients and risk scores (RS) provide insights.^{8,9}

Black-box approaches

- RF and XGB predict by capturing complex patterns.
- RF constructs multiple DTs on a training set.^{10,11}
- When classification occurs, predictors are aggregated across several trees.
- XGB optimizes an objective function using gradient descent to refine DTs.¹²
- It sequentially builds models to correct errors of previous ones.

Model building

- Model building considers splitting treatment decision-making into two steps:
 - Predicting whether a patient would switch therapy
 - Predicting which class of therapy the patient switches to
- Model performance was assessed using accuracy, area under the curve (AUC) receiver operating characteristic (ROC), AUC-precision recall (PR), expected calibration error (ECE) and Brier Score.
- AUC (max 1) considers model performance (both patients that switch and stay).
- High AUC-PR indicates the model can identify true positives among positive predictions.

Model fitting

- Overall, the dataset was split into training/model performance-validation (84%, 16%, 20%).
- All models were trained using the same data splits; parameters of each model were optimized through 10-fold cross-validation and random search.

Results (cont.)

Interpretable ML models

- These explained the attributes associated with switching therapy through models trained to predict treatment switching.
- DTs were most ideal, categorizing patients into non-overlapping groups and allowing for clinical decision-making.
- Decision sets (RE, BRS):
 - These performed best among the interpretable ML models, but had the highest calibration errors
 - They presented a set of rules ranked by importance, which indicates the contribution of linear items and rule items to the prediction of whether to switch treatment.
- Classification groups overlapped.

Table 2. RE model (excerpt) presents a set of rules ranked by importance

Rule	Coefficient	Importance*
History of btsDMARDs	2.850	1.141
History of btsDMARDs AND 20.7 ≤ CDAI ≤ 76 AND previous therapy csDMARDs	-0.416	0.205
History of btsDMARDs AND 20.7 ≤ CDAI ≤ 76 AND switched at last visit	-0.335	0.164

Abbreviations: bts = biologic-targeted synthetic; CDAI = Clinical Disease Activity Index; cs = conventional, synthetic; DMARD = disease-modifying anti-rheumatic drug; RE = rule ensemble.
*Indicates the contribution of linear items and rule items to the prediction of whether to switch treatment.

Decision lists (RL):

- These presented a list of if/then rules ranked by probability of switching therapy.

Table 3. RL model (excerpt) presents a list of rules ranked by the probability of switching therapy

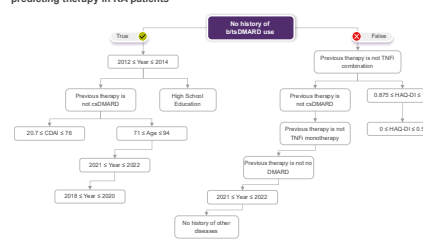
Conditions	Probability of Switching Therapy
If history of btsDMARDs	100%
If switch at last visit	42.4%
If 20.7 ≤ CDAI ≤ 76	34.6%

Abbreviations: bts = biologic-targeted synthetic; CDAI = Clinical Disease Activity Index.
*Indicates the contribution of linear items and rule items to the prediction of whether to switch treatment.

Decision trees (RT,DT):

- These presented series of decisions with final nodes.
- Patients were categorized into non-overlapping groups with clear decision paths.

Figure 1. DT of the therapy model (excerpt) in the combined model (DT+DT) for predicting therapy in RA patients



Abbreviations: bts = biologic-targeted synthetic; CDAI = Clinical Disease Activity Index; DMARD = disease-modifying anti-rheumatic drug; HAQ-DI = Health Assessment Questionnaire – Disability Index.

Linear models (LR, RS):

- LR showed the importance of weights of features by providing the top 5 features with the highest positive coefficients and the top 5 features with the largest absolute negative coefficients.
- Using RS, scores were constructed from four key variables:
 - Disease activity within specific range at baseline visit (1 point)
 - Switched therapy at last visit (1 point)
 - Previous therapy was combination therapy (-1 point)
 - History of biologic therapy (-5 points)
- RS presented scores of whether a patient will switch based on risk percentages.


Table 4. Using RS to predict whether patients will switch therapy

Score	0	1	2	3	4	5
Risk	11.5%	25.9%	50%	73.1%	95.9%	99.2%

Conclusions

- Interpretable ML models had comparable predictive properties to black-box models.
- Among interpretable ML models, DTs are ideal for describing switching patterns in RA since they can categorize patients into non-overlapping groups.
- Interpretable ML models offer transparency and understandability in the decision-making process, which is particularly important for a clinical audience.

Presented at the 41st International Society for Pharmacoepidemiology (ISPE) Annual Meeting | 22–26 August 2025 | Washington, DC, USA



Poster Title

Post-Approval Real-World Studies of RNA Therapies in the EU: What Can We Learn From Study Characteristics?

Objective

- To provide an overview of key design characteristics of studies that are registered in the Heads of Medicines Agencies (HMA)-EMA Catalogues (HMA-EMA)⁴ and ClinicalTrials.gov (CTgov)⁵ of real-world data (RWD) studies on RNA therapies approved in the EU.

<p>AUTHORS</p> <p>Sara Angleman¹ Elizabeth Donahue² Mai Duong³</p>	<p>AFFILIATION</p> <p>Nelson Ndegwa Gichora¹ Nahila Justo¹ Alice Rouleau⁴</p>	<p>PRACTICE AREA</p> <p>PPD™ Observational Studies, Thermo Fisher Scientific</p> <p>PPD™ Evidera™ Real-World Data & Scientific Solutions, Thermo Fisher Scientific at the time this study was completed</p>	<p>CLIENT</p> <p>INTERNAL</p>
--	---	--	--------------------------------------

Post-Approval Real-World Studies of RNA Therapies in the EU: What Can We Learn From Study Characteristics?



SLA-049

Sara Angleman¹, Elizabeth Donahue², Mai Duong³, Nelson Ndegwa Gichora¹, Nahila Justo¹, Alice Rouleau⁴
¹Thermo Fisher Scientific, Stockholm, Sweden; ²Thermo Fisher Scientific, Waltham, MA, USA; ³Thermo Fisher Scientific, London, UK; ⁴Thermo Fisher Scientific, Paris, France

Background

- Human gene therapy products include RNA therapies¹.
- RNA therapies are governed by the EMA (2010)¹ and FDA (2020)² guidance documents for gene therapy.
- According to the American Society of Gene and Cell Therapy, as of Q3 2024, 34 RNA therapies had been approved globally³.
- Of these, 10 were approved in the EU (RNA vaccines excluded from this abstract).
- The landscape of RNA therapies is developing quickly, and it is important to understand how best to design post-approval real-world studies of these therapies.

Objectives

- To provide an overview of key design characteristics of studies that are registered in the Heads of Medicines Agencies (HMA)-EMA Catalogues (HMA-EMA)⁴ and ClinicalTrials.gov (CTgov)⁵ of real-world data (RWD) studies on RNA therapies approved in the EU.

Methods

- A targeted review of the HMA-EMA Catalogues of RWD Sources and Studies (HMA-EMA)⁴ and ClinicalTrials.gov (CTgov)⁵ was conducted to identify RWD studies of approved RNA therapies. We extracted study design characteristics from each of 53 studies (Table 1).
- Records were retrieved from the databases, de-duplicated, and visually reviewed for any additional duplicates. All studies were individually reviewed to ensure they met the search criteria and extracted key information.
- Selection criteria in HMA-EMA:
 - No medical procedures
 - No health conditions
 - No clinical trials
 - No ecological studies
 - No case series
 - No self-control case series
- Selection criteria in ClinicalTrials.gov:
 - Excluded the expansion of access studies
 - SDU was defined as retrospectively collected data and data collected from medical records.
 - When the duration of the study was calculated from start and end dates, the number was rounded up to a whole number for the categorical duration of the study.
 - The comparator group was marked as "No" where there was no information available or if no comparison analysis was mentioned.

Table 1. Approved RNA Therapies

Product Name	Generic Name	Year first approved	Indication(s)	Locations approved by EMA/CEMS	Sponsor(s)	Number of RWE/S studies
Amviva	Vismodegvir	2022	Transferrin-related Hereditary Amyloidosis	US, Australia, UK, and EU Member Countries	Amylum Pharmaceuticals	3
Givlaari	givosiran	2020	Porphyria	US, Brazil, Japan, Canada, UK, and EU Member Countries	Amylum Pharmaceuticals	3
Leqvio	inclisiran	2020	Atherosclerosis, heterozygous familial hypercholesterolemia, hypertriglyceridemia	UK, US, China, Japan, South Korea, and EU Member Countries	Novartis	16
Nulectra	bedempipinil	2021	Methylglucuronidase Deficiency	US, UK, Israel, and EU Member Countries	Bentylol Therapeutics	1
Onpattro	patisiran	2018	Transferrin-related Hereditary Amyloidosis	Switzerland, Brazil, Australia, UK, and EU Member Countries	Amylum Pharmaceuticals	7
Orlumo	lumasiran	2020	Hyperoxaluria	US, Canada, Switzerland, and EU Member Countries	Amylum Pharmaceuticals	2
Qalody	tofersen	2023	Anyomitic Lateral Sclerosis	US, UK, Canada, and EU Member Countries	Biogen	1*
Spiraxza	nusinersen	2016	Spinal Muscular Atrophy	US, Canada, Japan, Switzerland, China, Brazil, and EU Member Countries	Biogen	17
Tegsedi	incisiran	2018	Transferrin-related Hereditary Amyloidosis	US, Canada, Italy, and EU Member Countries	Alice Therapeutics	3
Waylivra	volanesronam	2019	Hypotriglyceridemia, Lipoprotein Lipase Deficiency	Brazil, Argentina, Mexico, Chile, Colombia, Peru, Scotland, and EU Member Countries	Alice Therapeutics	1

Abbreviations: RWD = real-world data; RWE = real-world evidence
 *Not included in total study count of RWE/S studies (n=53) as there were no study details available for this therapy.

Results

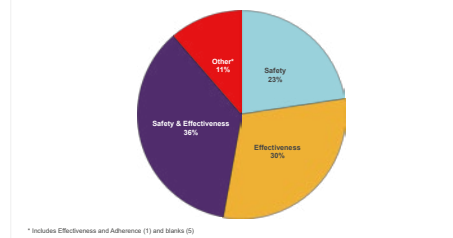
- 73 studies were identified in CTgov⁵ and 10 in the HMA-EMA⁴. After applying exclusion criteria and de-duplication, data on 53 studies were analyzed.
- Of the 53 studies identified, the study start date for 45 (85%) was after 2020, and for the remaining eight studies (2017–2019), it was before 2020. Median study duration (calculated from reported actual/planned start date until planned final study report date) was 2 years, with a range of <2–11 years. Nine studies included North American countries, one was in Latin America, 23 were in EMA, six were in Asia-Pacific (APAC), and 14 included multiple countries across more than one region (Table 2).
- 12 studies were planned to assess safety only, and 16 studies were planned to assess efficacy only; 19 studies planned to assess both safety and effectiveness, and another also planned to assess adherence in addition to efficacy. Five studies did not specify safety or effectiveness in their primary or secondary objectives (Figure 1). The study entries rarely specified the planned duration of safety follow-up per patient. Six studies were pregnancy safety studies. The indications were rare diseases for six of the 10 RNA therapies.
- 12 studies indicated a comparator cohort. The data source was difficult to discern from the information provided in the structured data source field: 11 studies reported data sources including prospective patient-based data collection, physician-completed CRFs, secondary existing data sources, EMR, and medical record review. The median target size for the studies was 150 patients, with a range of 10 to over 2000 patients. The planned study ages (number of studies) were adults (26), pediatric (2), and both adults and pediatric (24), and one study was missing age (Table 2).

Results (cont.)

Table 2. Study Characteristics

Variable	Number of Studies
Countries	
NA	9
EMEA	23
Asia-Pacific	6
Latin America	1
Multi-Regional	14
Duration	
>2 years	22
2–5 years	19
6–9 years	6
10+	6
Target Sample Size	
<50	13
50–100	10
101–200	9
201–500	7
501–1000	5
1001+	7
Blank	2
Age Range	
Pediatrics	2
Adults	26
Both	24
Blank	1
Started	
Before 2020	8 (2017–2019)
After 2020	45 (2020–current)
Completed	22
Ongoing	31
Objective	
Safety	12
Effectiveness	16
Both	19
Other	6

Figure 1. Post-Approval Study Objectives



Conclusions

- A notable increase in the number of real-world studies investigating RNA therapies has been observed since the release of the FDA's Gene Therapy Long-Term Follow-Up Guidance in 2020.
- The majority of post-2020 studies reflect a growing interest in real-world data collection for RNA therapies, particularly in rare diseases.
- There is a trend toward global, multi-regional studies, but substantial variation exists in study design, duration, and patient population.
- The average duration of post-approval studies is approximately four years. Although study registries do not consistently specify the rationale for study length, longer durations (ranging from 5–15 years) may be associated with the characteristics of the therapeutic delivery vector.
- Many studies lack clear specification of follow-up parameters, especially for safety.
- Data source transparency and uniformity remain limited, complicating cross-study comparison.
- Study designs are diverse, with a balance of safety and efficacy objectives, though some studies do not clearly define outcomes.

References

- EMA. Guidelines on Follow-Up of Patients Administered With Gene Therapy Medical Products. 2009. https://www.ema.europa.eu/en/documents/scientific-guideline/guideline_follow-patients-administered-gene-therapy-medical-products_en.pdf
- US FDA. Long Term Follow-up After Administration of Human Gene Therapy Products. 2020. <https://www.fda.gov/regulatory-information/search-fda-guidance-documents/long-term-follow-up-after-administration-human-gene-therapy-products>
- American Society of Gene + Cell Therapy. Gene, Cell, & RNA Therapy Landscape Report. 2024. <https://www.asgct.org/gene-cell-therapy-landscape-report-2024-report.aspx>
- EMA. Catalogue of RWD Studies. 2025. <https://catalogue.ema.europa.eu/catalogue-rwd-studies>
- ClinicalTrials.gov. 2025. <https://clinicaltrials.gov/>

Disclosures
 SA, ED, and AR are employees of PPD's Observational Studies, Thermo Fisher Scientific. MD, NNG, and NA are current employees of PPD. Other PPD Real-World Data & Scientific Solutions, Thermo Fisher Scientific, at the time this study was completed. Funding provided by Thermo Fisher Scientific.



ISPE 2025

22–26 August 2025 | Washington, DC, USA

Poster Title

Beyond Comparison: Target Trial Emulation (TTE) Principles Strengthen Observational Single-arm (SA) Effectiveness Studies

Objective

- Describe implementation of TTE principles for SA effectiveness studies based on registry data.

AUTHORS Robert R. McLean Oksana Pugach Melissa Moore-Clingenpeel Adam P. Sima Leslie R. Harrold Heather J. Litman	AFFILIATION Thermo Fisher Scientific, Waltham, MA, USA	PRACTICE AREA PPD™ CorEvitas™ Clinical Registries, Thermo Fisher Scientific	CLIENT INTERNAL
--	--	---	---------------------------



Beyond Comparison: Target Trial Emulation (TTE) Principles Strengthen Observational Single-arm (SA) Effectiveness Studies

Robert R. McLean¹, Oksana Pugach¹, Melissa Moore-Clingenpeel¹, Adam P. Sima¹, Leslie R. Harrold¹, Heather J. Litman¹
¹ Thermo Fisher Scientific, Waltham, MA, USA

Background

- Single-arm (SA) studies, where all participants receive the same treatment, are used across the drug development cycle to provide evidence supporting a drug's value story.¹
- In the post-marketing space, observational SA studies can answer questions about drug utilization in patient care (e.g., treatment patterns) and clinical impact in real-world settings.
- A common objective is to demonstrate effectiveness in more broad patient populations than those included in pivotal trials, but in a non-comparative manner, i.e., no reference group.
- Patient registries are valuable for addressing this question as they provide an efficient means of describing short- and long-term effectiveness by collecting clinician- and patient-reported outcomes (ClinROs/PROs) not available in other data sources, e.g., claims, electronic health records.
- Although SA effectiveness studies using registry data do not support causal inference, leveraging Target Trial Emulation (TTE) principles^{2,3} can help mitigate threats to validity and strengthen conclusions.

Objective

- Describe implementation of TTE principles for SA effectiveness studies based on registry data.

Methods

PPD™ CorEvitas™ Clinical Registries

Registry Design

- General design across therapeutic areas – dermatology, rheumatology, gastroenterology, neurology.
- Prospective, non-interventonal (observational) registries of patients diagnosed with condition, under the care of a healthcare provider.
- Longitudinal, regulatory-grade data collected from patients and providers via questionnaires administered during routine clinical encounters occurring every ~6 months.

Eligibility Criteria

- Criteria vary by therapeutic area but generally include:
 - Diagnosed with the condition by a specialist.
 - At least 18 years of age (adolescent registries excepted).
 - Willing and able to provide written consent for participation in the registry.
 - Willing and able to provide personally identifiable information, including, at a minimum, full name and date of birth.
 - Receiving a new (never used) enrollment eligible medication, started within 12 months prior to the registry enrollment visit or prescribed at the enrollment visit.
 - Eligible medications typically comprise approved biologic or targeted synthetic medications (advanced therapies) for treatment of the specified condition.
 - Those participating or planning to participate in a double-blind randomized trial are not eligible for enrollment.

Data Elements

- Demographics – age, sex, race/ethnicity, marital status, education, employment, insurance.
- Lifestyle variables – body mass index, alcohol use, smoking.
- Medical history – comorbidities, clinical events, infections, vaccinations.
- Disease characteristics – disease onset/duration, symptoms.
- Medications – eligible medication start/stop dates and associated reasons (including history prior to enrollment), dose/frequency, other concomitant medications.
- ClinROs/PROs – commonly utilized in clinical trials.
- Adverse events – includes validated serious and targeted safety events.

TTE Elements

- Eligibility
- Treatment Strategy
- Assignment Procedure
- Follow-up Period
- Outcome
- Causal Contrast
- Analysis

Applying TTE Elements to SA Effectiveness Studies

- Examples are described for how each TTE element may be applied for strengthening SA effectiveness studies of immunomodulator medications.
- Applications are based on leveraging data from the PPD CorEvitas Clinical Registries.

Disclosures

RRR, OP, MRG, AP, LM, and MLC are employees of PPD™ CorEvitas™ Clinical Registries, who received no external funding to conduct this study. This study is sponsored by Thermo Fisher Scientific. PPD CorEvitas Clinical Registries has been supported through contractual agreements in the last two years by AbbVie, Amgen, Inc., AstraZeneca, Biogen, Bristol Myers Squibb, Celgene, Eisai, Inc., Genentech, Genentech, Inc., Janssen Pharmaceutica, Inc., Lilly, Pfizer, Novartis, Celis, Celis Therapeutics, Pfizer, Inc., Sanofi, Pharmaceutical Research and Manufacturers of America (PhRMA).

Acknowledgments

Editorial and graphic design support were provided by Michael Gross and Brent Berger at Thermo Fisher Scientific.

Results

- Eligibility**
 - Comprehensive data capture on patient characteristics, disease activity, and medication use (e.g., start dates, prior therapy use) facilitates:
 - Defining the study cohort using information up to time zero and not beyond.
 - Tailoring the cohort to specific populations of interest, e.g., including all patients prescribed the drug or restricting to those with baseline moderate to severe disease.
- Treatment Strategy**
 - Data ascertained by the registry including concomitant therapies, drug dose and frequency, and discontinuations and switches (as well as their reasons) may be used to specify relevant treatment protocol and identify deviations.
- Assignment Procedure**
 - SA effectiveness studies are non-comparative by design with all individuals receiving the same drug; thus, emulation of random assignment to treatment is not applicable.
- Follow-up Period**
 - Particularly in immunology/neurology settings, clinically meaningful effectiveness typically focuses on response at a specific timepoint (e.g., 6 months).
 - Because PPD CorEvitas Registries are observational and clinic visits may not occur at a consistent cadence, follow-up can be based on visits occurring within windows around the timepoint of interest (e.g., 5–9 months for a 6-month follow-up).
- Outcome**
 - ClinRO/PRO data includes validated measures used in randomized trials and may be used to clearly define outcomes, typically as a binary response measure or change from time zero at the specified timepoint.
 - If safety is of interest, registry data obviates reliance on complex algorithms based on Current Procedural Terminology or International Classification of Diseases codes to identify adverse events, as is necessary in claims databases.
- Causal Contrast**
 - While there are no causal contrasts in a SA study, the approach may be specified to address the research question of interest, e.g., evaluate all follow-up time, regardless of treatment protocol deviations (intention-to-treat) vs. evaluate follow-up time under which a patient is adherent to treatment protocol (per protocol).
- Analysis**
 - Descriptive measures ubiquitous in SA studies.
 - Analyses may be stratified by prior treatment (e.g., naive vs. experienced, line of therapy).
 - The analytic plan should include descriptions of how missing data and loss to follow-up (e.g., net reclassification improvement, inverse probability of censoring weighting) or time-varying confounding (e.g., inverse probability of treatment weighting, marginal structural model) are handled.

Limitations

- Because SA effectiveness studies have no comparator group, conclusions should be considered carefully.
 - Causal inferences are prohibited, and any conclusions should be contextualized based on external information from relevant studies.
 - Findings cannot be considered as conclusive evidence of drug effectiveness.

Conclusion

- When comprehensive, granular, regulatory-grade registry data are available, TTE principles can be applied to SA effectiveness studies to produce robust, clinically informative evidence that provides contextualization in the absence of other real-world data sources.

References

- Opejewa O, Dargatzis D, Bhatnagar C, Bhatnagar N, et al. *Pharm Stat*. 2024;19(1):1296-1306-1317.
- Hernan MA, Rubin JB. *Am J Epidemiol*. 2016;183(10):758-64.
- Zhou SB, Liu H, Solomon DH, Yoshida K. *Ann Rheum Dis*. 2020;39(7):885-892.

ppd

Presented at the 41st International Society for Pharmacoepidemiology (ISPE) Annual Meeting | 22–26 August 2025 | Washington, DC, USA

ISPE 2025

22–26 August 2025 | Washington, DC, USA

Poster Title

Longitudinal, Regulatory-grade Registry Data Can Be Used to Emulate a Target Trial: An Application on Evaluating Treatments for Plaque Psoriasis

Objective

- To demonstrate how registry data can fit a TTE framework to answer comparative effectiveness and safety questions about real-world medication use.

AUTHORS Heather J. Litman ¹ Robert R. McLean ¹ Oksana Pugach ¹ Adam Sima ¹ Melissa Moore-Clingenpeel ¹ Leslie Harrold ¹	AFFILIATION ¹ Thermo Fisher Scientific, Waltham, MA, USA ² Chalmers University and University of Gothenburg, Gothenburg, Sweden	PRACTICE AREA PPD™ CorEvitas™ Clinical Registries, Thermo Fisher Scientific	CLIENT INTERNAL Chalmers University and University of Gothenburg
--	--	---	--

Longitudinal, Regulatory-grade Registry Data Can Be Used to Emulate a Target Trial: An Application on Evaluating Treatments for Plaque Psoriasis

Thermo Fisher Scientific C-149

Heather J. Litman¹, Robert R. McLean¹, Oksana Pugach¹, Adam Sima¹, Melissa Moore-Clingenpeel¹, Leslie Harrold¹
¹ Thermo Fisher Scientific, Waltham, MA, USA

Background

- After a medication is approved, real-world data is essential and necessary to complement randomized controlled trials (RCTs).
- RCT populations are selected according to narrow inclusion/exclusion criteria and treated in a highly controlled environment, but real-world data (RWD):
 - Feature outcomes in routine clinical care, better reflecting the complexity and diversity of clinical practice.
 - Provide a more nuanced understanding that informs healthcare decision-making.
 - Help identify a well-defined market and strategy for reaching patients who will benefit most from a new drug.
- Comparative effectiveness and safety studies of RWD can apply the target trial emulation (TTE) framework to address causal inferences about the benefits and risks of a medication.¹⁻³
- The introduction of the first biologic agents in psoriasis (PsO) more than 10 years ago was followed by a wave of highly effective medications.
 - Within PsO, RWD from registries within a TTE framework have been used to answer comparative effectiveness and safety questions to guide dermatologists in treatment decision-making for their patients.

Objective

- To demonstrate how registry data can fit a TTE framework to answer comparative effectiveness and safety questions about real-world medication use.

Methods

TTE Framework

- Extends causal inference principals rooted in RCTs.
- Observational data typically emulates pragmatic trials (treatments compared under usual conditions to which they're applied) vs. RCTs (tight monitoring, enforcement of adherence, blinding).
- Follows parameters of a hypothetical trial, addressing bias through a transparent structured study design approach with two steps:
 - Defining the target trial protocol
 - Emulating that protocol using observational data

Focus on Seven Key TTE Principles

- Eligibility criteria
- Treatment strategies
- Treatment procedures
- Follow-up period
- Outcomes
- Causal contrasts
- Analysis plan

PPD™ CorEvitas™ PsO registry

- Prospective, non-interventional multicenter registry.
- Collects data on adults diagnosed with PsO by dermatologists and receiving a biologic or targeted synthetic drug for treatment of PsO.
- Baseline and follow-up data on patient characteristics, drug use, disease characteristics, and clinical and patient-reported outcomes collected from patients and providers during routine clinical encounters occurring at 6-month intervals.

Disclosures

All authors are employees of PPD™ CorEvitas™ Clinical Registries and subsidiaries of Thermo Fisher Scientific, and did not receive outside funding to conduct this study. This study is sponsored by Thermo Fisher Scientific. CorEvitas Clinical Registries has been registered by controlled submissions in the last 2 years by AbbVie, Amgen, Inc., Astra, Biologics, Regeneron, Bristol Myers Squibb, Cytos, Eli Lilly and Company, Genentech, GSK, Janssen Pharmaceutica, Inc., LEO Pharma, Novartis, Otsuka Pharmaceutical, Pfizer, Inc., Sun Pharmaceutical Industries Ltd., and UCB S.A.

Acknowledgments

Editorial and graphic design support were provided by Michael Givoni and Shari Berger of Thermo Fisher Scientific.

Results

To emulate a target trial for comparing treatments for plaque PsO:
Eligibility criteria, treatment strategies, and treatment procedures

- Identify at baseline, defined as the registry visit when the medication was started ("time zero").

- Eligibility criteria**
 - Collect registry information on patient characteristics (e.g., age, comorbidities), disease activity, and past and current medication use at "time zero."
 - Include patients initiating treatment with a drug of interest or comparator who are followed from treatment initiation.
 - Use past behavior to determine who is likely to have follow-up visits.
- Treatment strategies**
 - In the PsO treatment protocol definition, include specific drug initiated, prior drug use, dosage, frequency, and concomitant medications at "time zero."
 - Update treatment status over follow-up, including start/stop dates, changes in dose/frequency and concomitant medications, and reasons for stops, to identify specified protocol violations (e.g., drug stops due to economic reasons, starting comparator drug or another drug other than the treatment of interest).
 - Consider initiators of the treatment of interest or a comparator at "time zero."
- Treatment procedures**
 - Assign patients to treatment group for analysis based on their first eligible treatment initiation.
 - Consider an active comparator to improve emulation of random assignment.
 - Capture relevant baseline characteristics in a robust manner to facilitate adequate pseudo-randomization.
 - Use stabilized propensity score weights (inverse probability of treatment weighting) to adjust for measured confounding at "time zero," emulating randomization at baseline.
- Follow-up period**
 - Because clinically meaningful short-term effectiveness in clinical trials focuses on response at a specific timepoint, consider using a specific follow-up visit (e.g., 6 months).
 - Since the PsO registry is observational and clinic visits may not occur precisely at 6 months, use a window around the timepoint (e.g., 5-9 months).
 - Align treatment assignment, study eligibility, & start of outcome tracking at "time zero."
- Outcomes**
 - Use validated clinician- and patient-reported measures, similar to RCTs.
 - Consider change from "time zero" in disease activity for PsO.
- Causal contrasts**
 - Consider two options depending on how non-adherence (deviation from treatment strategy) was handled:
 - Intention-to-treat contrast (preferred approach):
 - Include all initiators of treatment with deviations from treatment strategy ignored.
 - Represents real-world effect of being assigned to treatment.
 - Per-protocol contrast:
 - Retain non-adherent subjects and adjust using analytic methods.
 - Represents effect of the intended treatment strategy that would have been observed if all subjects had adhered to their intended treatment strategy.
- Analysis plan**
 - Design and conduct analysis informed by treatment strategies and causal contrast.
 - Include how missing data/list-to-follow-up will be treated in the analysis.

Conclusions

- Longitudinal observational registry data such as the CorEvitas PsO Registry can be used to design strong studies for causal inference within the TTE framework.

References

- DukeMargolis Institute for Health Policy. "Real-World Evidence to Support Causal Inference: Methodological Considerations for Non-Interventional Studies." June 2024.
- Hernán MA, Royston JP. *Am J Epidemiol*. 2016;183(8):758-64.
- Zhao SB, et al. *Ann Rheum Dis*. 2022;79(7):885-890.

Presented at the 41st International Society for Pharmacoepidemiology (ISPE) Annual Meeting | 22–26 August 2025 | Washington, DC, USA

© 2025 Thermo Fisher Scientific, Inc. All rights reserved. All trademarks are the property of Thermo Fisher Scientific unless otherwise indicated. PPD™ is a trademark of PPD. The other trademarks are the property of their respective owners. PPD and CorEvitas are trademarks of PPD.

ppd




Poster Title

Policy Impact on Pregnancy Registry Participation and Reporting of Maternal/Infant Outcomes

Objective

- To review published and/or gray literature, defined as information produced outside of academic or commercial publishing pathways (e.g., technical reports, white papers) to assess the potential impact of the Dobbs ruling on:
 - Objective 1: Pregnancy registry enrollment, loss to follow-up, and missing data.
 - Objective 2: Maternal and infant outcomes, particularly in congenital anomalies, which is a critical outcome studied in pregnancy registries.

<p>AUTHORS</p> <p>Ronna L. Chan, PhD, MPH Syd Phillips, MPH Kathleen Shannon, MPH Yao Yao, MS</p>	<p>AFFILIATION</p> <p>Thermo Fisher Scientific, Waltham, MA, USA</p>	<p>PRACTICE AREA</p> <p>PPD™ Observational Studies, Thermo Fisher Scientific</p>	<p>CLIENT</p> <p>INTERNAL</p>
--	---	---	--------------------------------------



B-265

Policy Impact on Pregnancy Registry Participation and Reporting of Maternal/Infant Outcomes

Ronna L. Chan, PhD, MPH; Syd Phillips, MPH; Kathleen Shannon, MPH; Yao Yao, MS
Thermo Fisher Scientific, Waltham, MA, USA

Background

- Prospective pregnancy registries monitor the safety of exposure of newly approved medicinal products on maternal and infant outcomes.
- The success of the registries in the evaluation of the association between maternal exposure to medicinal products and outcomes of mothers and infants depends on the willingness of pregnant individuals to enroll and provide detailed information such as:
 - Reproductive and pregnancy history
 - Maternal and pregnancy complications
 - Pregnancy outcome
- However, due to the US Supreme Court's decision in *Dobbs v. Jackson Women's Health Organization* (a.k.a. the *Dobbs* ruling) in June 2022,¹ this has resulted in reduced access to legal and safe induced abortion across many US states.
- Consequently, changes in registry enrollment, loss to follow-up, and missing maternal and infant data are anticipated.

Objectives

- To review published and/or gray literature, defined as information produced outside of academic or commercial publishing pathways (e.g., technical reports, white papers) to assess the potential impact of the *Dobbs* ruling on:
 - Objective 1:** Pregnancy registry enrollment, loss to follow-up, and missing data.
 - Objective 2:** Maternal and infant outcomes, particularly in congenital anomalies, which is a critical outcome studied in pregnancy registries.

Methods

- We performed targeted literature reviews of published and gray literature.

Objective 1

We used variations on the following search terms in published literature: abortion, *Dobbs*, pregnancy, and prospective registry.

Objective 2

We used variations on the following search terms in published and gray literature: abortion, pregnancy termination, maternal health outcome, infant health outcome, and congenital abnormalities/anomalies.

- The authors reviewed articles that were identified using the search terms to determine relevancy.

Results

Pregnancy Registry Enrollment, Loss to Follow-Up, and Missing Data

- The targeted literature review from the published literature identified four articles and one³ contained relevant information.
- In the identified article,² the researchers observed a 27.1% reduction in potential enrollment in registries in states with banned or restricted abortion rights comparing the post-versus pre-*Dobbs* periods.
- Table 1** presents published literature search terms and results for Objective 1.

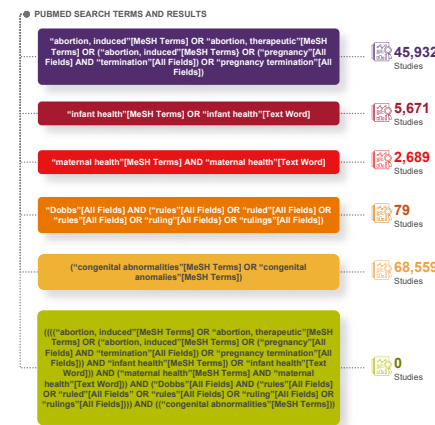
Step	PubMed Search Terms	Results
1	"abortion, induced"[MeSH Terms] OR "abortion, therapeutic"[MeSH Terms] OR "abortion, legal"[MeSH Terms] OR "abortion, illegal"[MeSH Terms]	44,760
2	"abortion"[Text Word]	101,392
3	"induced abortion"[Text Word]	6,623
4	"therapeutic abortion"[Text Word]	1,689
5	"legal abortion"[Text Word]	1,327
6	"illegal abortion"[Text Word]	420
7	"abortion, induced"[MeSH Terms] OR "abortion, therapeutic"[MeSH Terms] OR "abortion, legal"[MeSH Terms] OR "abortion, illegal"[MeSH Terms] OR "abortion"[Text Word] OR "induced abortion"[Text Word] OR "therapeutic abortion"[Text Word] OR "legal abortion"[Text Word] OR "illegal abortion"[Text Word]	102,132
8	"Dobbs"[Text Word]	539
9	"Dobbs ruling"[Text Word]	29
10	"Dobbs"[Text Word] OR "Dobbs ruling"[Text Word]	529
11	"registry"[Text Word]	177,590
12	"pregnancy registry"[Text Word]	398
13	"prospective study"[Text Word]	1,74,442
14	"Prospective Studies"[MeSH Terms]	707,881
15	"prospective"[Text Word]	1,046,497
16	"registry"[Text Word] OR "pregnancy registry"[Text Word] OR "prospective study"[Text Word] OR "Prospective Studies"[MeSH Terms] OR "prospective"[Text Word]	1,195,201
17	"abortion, induced"[MeSH Terms] OR "abortion, therapeutic"[MeSH Terms] OR "abortion, legal"[MeSH Terms] OR "abortion, illegal"[MeSH Terms] OR "abortion"[Text Word] OR "induced abortion"[Text Word] OR "therapeutic abortion"[Text Word] OR "legal abortion"[Text Word] OR "illegal abortion"[Text Word] AND ("Dobbs"[Text Word] OR "Dobbs ruling"[Text Word]) AND ("registry"[Text Word] OR "pregnancy registry"[Text Word] OR "prospective study"[Text Word] OR "Prospective Studies"[MeSH Terms] OR "prospective"[Text Word]) AND ("Pregnancy"[Text Word] OR "Pregnancy"[MeSH Terms])	4

Results (cont.)

Maternal and Infant Outcomes

- The targeted literature review from published and gray literature identified one article.³
- Figure 1** presents the published literature search for Objective 2.

Figure 1. Published Literature Search for Objective 2



- "abortion, induced"[MeSH Terms] OR "abortion, therapeutic"[MeSH Terms] OR "abortion, induced"[MeSH Terms] OR ("pregnancy"[All Fields] AND "termination"[All Fields]) OR "pregnancy termination"[All Fields]
- "infant health"[MeSH Terms] OR "infant health"[Text Word]
- "maternal health"[MeSH Terms] AND "maternal health"[Text Word]
- "Dobbs"[All Fields] AND ("rules"[All Fields] OR "ruled"[All Fields] OR "rules"[All Fields] OR "ruling"[All Fields] OR "rulings"[All Fields])
- ("congenital abnormalities"[MeSH Terms] OR "congenital anomalies"[MeSH Terms])
- (("abortion, induced"[MeSH Terms] OR "abortion, therapeutic"[MeSH Terms] OR "abortion, induced"[MeSH Terms] OR ("pregnancy"[All Fields] AND "termination"[All Fields]) OR "pregnancy termination"[All Fields]) AND "infant health"[MeSH Terms] OR "infant health"[Text Word]) AND ("maternal health"[MeSH Terms] AND "maternal health"[Text Word]) AND ("Dobbs"[All Fields] AND ("rules"[All Fields] OR "ruled"[All Fields] OR "rules"[All Fields] OR "ruling"[All Fields] OR "rulings"[All Fields]) AND ("congenital abnormalities"[MeSH Terms] OR "congenital anomalies"[MeSH Terms]))

- Table 2** shows the published and gray literature web search for Objective 2.

Web Search Terms	Results
Dobbs, Dobbs ruling, abortion, pregnancy termination, infant health outcomes, maternal health outcomes, congenital anomalies	1

- Research by Singh and Gallo³ suggested the *Dobbs* decision may have an adverse effect on maternal, fetal, and infant outcomes. Using monthly national data between 2018 and 2023 from the Centers for Disease Control and Prevention, 204 excess death (10% absolute increase) in infant mortality with congenital anomalies were attributed to after *Dobbs* (June 2022 – December 2023) compared to the previous five years (January 2018 – May 2022).

Conclusions

- Results from Objective 1** targeted literature review suggested a potential decline in registry enrollment. Such reactions may precipitate loss to follow-up in states where safe/legal abortion is not available.
- Participants who enroll in a registry may be lost to follow-up if they obtain an abortion post-enrollment and they and/or their healthcare providers decline to provide follow-up data related to the induced abortion due to negative repercussions.
- A targeted literature review for Objective 2 suggested that the *Dobbs* ruling may result in an increase incidence of adverse maternal, infant, and child outcomes potentially due to the restricted access to abortion care. This is further supported by Gemmill et al.,⁴ who found that the abortion ban in Texas was associated with a 12.9% increase in infant deaths in 2022.
- More research is needed to fully understand the impact of the *Dobbs* ruling on pregnancy registry enrollment, loss to follow-up, and the maternal and infant outcomes.

References

- Dobbs v. Jackson Women's Health Organization, 597 U.S. __ (2022).
- Wills MD, et al. *Pediatr Perinat Epidemiol*. 2024 Sep;38(7):627-634.
- Singh P, Gallo AF. *JAMA Pediatr*. 2024;178(12):1364-1366.
- Gemmill A, et al. *JAMA Pediatr*. 2024;178(8):764-791.

Disclosures

R.L.C., S.P., and Y.Y. are current or former employees of PPD™ Observational Studies, Thermo Fisher Scientific. Funding provided by Thermo Fisher Scientific.

Acknowledgments

Editorial and graphic design support were provided by Kawthar Nakayima of Thermo Fisher Scientific.

Presented at the 41st International Society for Pharmacoeconomics (ISPE) Annual Meeting | 22–26 August 2025 | Washington, DC, USA

© 2025 Thermo Fisher Scientific, Inc. All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording, or by any information storage and retrieval system, without prior written permission from Thermo Fisher Scientific. PPD™ is a trademark of PPD™ Observational Studies, Thermo Fisher Scientific. All other trademarks are the property of their respective owners.


Poster Title

Nuances of Risk Window Selection for Evaluation of Acute Adverse Events

Objective

- To examine the implications of using multiple RCWs, with emphasis on eligibility of exposure episodes and indexing of events.

<p>AUTHORS</p> <p>Heather Chubb¹ Bryan Cherry¹ Nicole Foster¹ Christine Barr¹ Melissa Moore-Clingenpeel¹ Dimitrios Pappas¹ Nathan Rubin¹ Elsie L Grace²</p>	<p>AFFILIATION</p> <p>¹Thermo Fisher Scientific, Waltham, MA, USA ²Eli Lilly and Company, Indianapolis, IN, USA</p>	<p>PRACTICE AREA</p> <p>PPD™ CorEvitas™ Clinical Registries, Thermo Fisher Scientific</p>	<p>CLIENT</p> <p>INTERNAL Eli Lilly and Company</p>
---	---	--	---



C-156

Nuances of Risk Window Selection for Evaluation of Acute Adverse Events

Heather Chubb¹, Nicole Foster¹, Melissa Moore-Clingenpeel¹, Nathan Rubin¹, Bryan Cherry¹, Christine Barr¹, Dimitrios Pappas¹, and Elsie L Grace²
¹Thermo Fisher Scientific, Waltham, MA, USA; ²Eli Lilly and Company, Lilly Corporate Center, Indianapolis, IN, 46285, USA

Background

- Real-world safety studies must account for the time after medication discontinuation during which a drug may influence the risk of adverse events (AEs); this is defined as a risk continuation window (RCW).
- In practice, the choice of RCW is not straight forward, especially when treatment cohorts contain medications with differing half-lives.
- The RCW has implications beyond the additional time at risk, which are not always obvious and may affect study results.

Objectives

- To examine the implications of using multiple RCWs, with emphasis on eligibility of exposure episodes and indexing of events.

Methods

Design and Data Source

- This was a comparative cohort study using prospectively collected data from the PPD™ CorEvitas™ Psoriasis (PsO) Registry.
- Participants were in the PsO Registry, aged 18+ years, and starting a registry-qualifying PsO treatment at or after registry enrollment (April 2015–September 2024).

Exposure Cohorts

- Exposure cohorts were starting treatment with an interleukin (IL)-17A or non-IL-17A biologic medication for the treatment of PsO and had at least one day of follow-up or a reported AE.

Exposure Time Definition

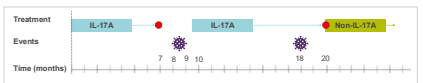
- Start of drug treatment until the earliest of:
 - Last administration or discontinuation (plus 30-, 60-, or 90-day RCW)
 - Start of new advanced therapy (other IL-17As, non-IL-17As, or Janus kinase inhibitors [JAKis])
 - End of study period
 - Loss to follow-up
 - Death
- Multiple exposure episodes could be created for one patient.
- Exposure episodes were assigned to the treatment cohorts using an "as-treated" approach accommodating medication restarts.

Defining Exposure Episodes and Mapping Events

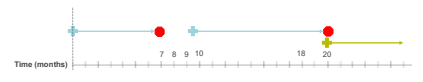
Example Study Outcome: Serious Infection (SI)

- The RCW under study (either 30-, 60-, or 90-days) represents time "at-risk" after the medication was stopped or temporarily disrupted.

Hypothetical patient treatment and event journey:

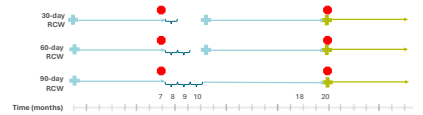


Step 1: Identify drug start and discontinuation dates



- A new index date (indicated by * or *) was assigned when a subject restarted or switched medication.

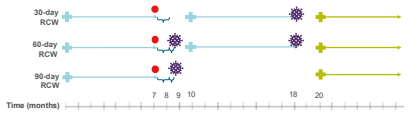
Step 2: Apply RCWs and combine exposure episodes as needed



- If a new advanced therapy (other IL-17A biologics, non-IL-17A biologics, or JAKis) was started during the RCW under study, starting the new medication stopped the RCW, and a new exposure episode began.
- Restarts of the same medication within the RCW were treated as continuous use (one episode).

Defining Exposure Episodes, Event Mapping (cont.)

Step 3: Map events onto exposure episodes



- Applying the 90-day RCW, the IL-17A restart within 90 days of medication discontinuation resulted in a single exposure episode.
- Only the first study outcome occurrence was reported within each RCW, so the second occurrence was not captured in the 90-day RCW.

Results

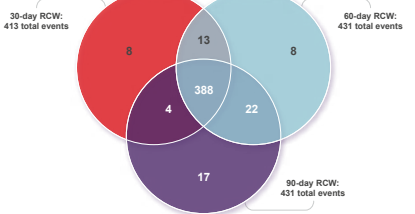
Table 1. Hypothetical patient

30-day RCW	60-day RCW	90-day RCW
3 episodes	3 episodes	2 episodes
1 IL-17A start	1 IL-17A start	1 IL-17A start
1 IL-17A restart	1 IL-17A restart	1 non-IL-17A start
1 non-IL-17A start	1 non-IL-17A start	

Abbreviations: IL = interleukin; RCW = risk continuation window

- The number of exposure episodes and number of events assessed varied by RCW.

Figure 1. Events analyzed across RCWs, propensity score matched study population



Abbreviation: RCW = risk continuation window

- Using different RCWs resulted in distinct exposure cohorts, each with varying numbers of eligible episodes, different matched populations, and different study outcomes mapped into exposure episodes.
- Results of the comparative analysis of IL-17A vs. non-IL-17A on risk of SI were similar across RCW definitions.
- However, cohorts based on different RCWs could not be directly compared due to having different populations.

Limitations

- We found little impact on overall study conclusions; this could be due to small numbers of events overall, the relative proportion of medication restarts vs. starts, the timing of the restarts, and consistent baseline characteristics across not only treatment cohorts but also risk window definitions.

Conclusions

- The RCW does not solely impact whether events following treatment discontinuation are included or excluded from exposure episodes; it also impacts the eligibility of subsequent exposure episodes and whether study outcomes are counted.
- Different index dates → different baseline characteristics → different analysis populations after PS matching
- This study ultimately demonstrated little impact of the risk-window definition on the comparative analysis of SIs; however, we illustrated ways in which modifications of the risk-window could have unintended consequences.
- Evaluating multiple RCWs in a single study may obfuscate findings and is inadvisable.

Disclosures and Acknowledgments

Funding for this poster was provided by Thermo Fisher Scientific. The PPD™ CorEvitas™ Clinical Registries have been supported through contracted subscriptions in the last few years by AbbVie, Amgen, Inc., Astra, Boehringer Ingelheim, Bristol Myers Squibb, Chicago, Eli Lilly and Company, Genentech, Janssen Pharmaceutica, Inc., LEO Pharma, Novartis, Ortho Dermatology, Pfizer, Inc., Sanofi Pharmaceutical, Sunovion, Inc. and UCB. The PsO Registry was developed in collaboration with the National Psoriasis Foundation (NPF). HC, MMC, CB, and DP are stockholders of Thermo Fisher Scientific; EG is a stockholder of Eli Lilly. Editorial and graphic design support was provided by Caroline Cole and Kaeleth Nakayama of Thermo Fisher Scientific.

Poster Title

Mandated Long-Term Follow-Up Studies of Patients Treated With Gene and Cell Therapies Approved in the European Union (EU): An Analysis of Commonalities and Variations

Objective

- The aim of this study was to describe characteristics and identify commonalities and variations of LTFU studies for the 16 GCTs approved in the EU until 31/03/2025.

Table 1 includes a list of therapies, the year of the first approval, indications, and other key information.

AUTHORS Nahila Justo ¹ Mai Duong ² Pingling Zeng ¹	AFFILIATION Elizabeth Donahue ³ Alice Rouleau ⁴ Sara Angleman ¹	PRACTICE AREA PPD™ Evidera™ Real-World Data & Scientific Solutions, Thermo Fisher Scientific PPD™ Observational Studies, Thermo Fisher Scientific	CLIENT INTERNAL
---	--	--	---------------------------

Mandated Long-Term Follow-Up Studies of Patients Treated With Gene and Cell Therapies Approved in the European Union (EU): An Analysis of Commonalities and Variations



Nahila Justo¹, Mai Duong², Pingling Zeng¹, Elizabeth Donahue³, Alice Rouleau⁴, Sara Angleman¹
¹ Thermo Fisher Scientific, Stockholm, Sweden; ² Thermo Fisher Scientific, London, UK; ³ Thermo Fisher Scientific, Waltham, MA, USA; ⁴ Thermo Fisher Scientific, Paris, France

Background

Regulators are intensifying efforts to reduce uncertainty around the long-term effects of gene and cell therapies (GCTs). Long-term follow-up (LTFU) of patients treated in clinical trials (CTs), imposed on investigational GCTs is often supplemented with mandated long-term follow-up post authorization (PA) studies.
The backbone of the rapidly evolving normative framework includes Regulation (EC) No 1394/2007, the EMA guideline on safety and efficacy follow-up and risk management of ATMPs, and the EMA's recently revised guideline on quality, non-clinical, and clinical requirements for investigational ATMPs in clinical trials.
GCT-specific guidance and norms interact with other critical compliance obligations, such as Good Pharmacovigilance Practices (GVP Modules VI, VII, VIII, X), the guideline on registry-based studies (EMA/426390/2021), the Scientific Guidance on Post-Authorisation Efficacy Studies (EMA/PDC/CAT/CMDH/FRAC/CHMP/261500/2015), and the guideline on safety and efficacy follow-up for advanced therapy medicinal products (EMA/414699/2020).

Objectives

The aim of this study was to describe characteristics and identify commonalities and variations of LTFU studies for the 16 GCTs approved in the EU until 31/03/2025. Table 1 includes a list of therapies, the year of the first approval, indications, and other key information.

Table 1. List of Gene and Cell Therapies Approved in the European Union As End of 2024 and the Associated LTFU Studies

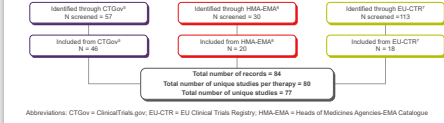
Product name	Year first approved	Indications	Therapeutic class	Sponsor(s)	Identified through	Source	Goal	Follow-up duration
Abema	2021	Multiple myeloma	Haematological cancers	Biosart AG (Brazil/Spain)	CTGov 2 ¹ EMA/EMA 1 ¹ EU/CTR 1 ¹	LTFU-CTP 7 LTFU-PA 3	Both 5	10-14 years 1
Beveq	2024	Hemophilia B	Haematology others	Pfizer	CTGov 2 ¹ EMA/EMA 1 ¹ EU/CTR 1 ¹	LTFU-CTP 2 LTFU-PA 2	Both 2	10-14 years 1
Beysfort	2021	Diffuse large B-cell lymphoma, follicular lymphoma, chronic lymphocytic leukemia, mantle cell lymphoma	Haematological cancers	Celgene (Brazil/Spain)	CTGov 2 ¹ EMA/EMA 1 ¹ EU/CTR 1 ¹	LTFU-CTP 1 LTFU-PA 1	Both 1	10-14 years 2
Caryfil	2022	Multiple myeloma	Haematological cancers	Legend Biotech (Japan/UK)	CTGov 2 ¹ EMA/EMA 2 ¹ EU/CTR 2 ¹	LTFU-CTP 4 LTFU-PA 3	Both 5	10-14 years 1
Cegary	2023	Stroke and ischemic stroke	Haematology others	CRISPR Therapeutics	CTGov 2 ¹ EMA/EMA 1 ¹ EU/CTR 1 ¹	LTFU-CTP 2 LTFU-PA 1	Both 3	10-14 years 3
Hemgenix	2022	Hemophilia B	Haematology others	uniQure (US, Belgium)	CTGov 2 ¹ EMA/EMA 1 ¹ EU/CTR 1 ¹	LTFU-CTP 4 LTFU-PA 2	Both 6	10-14 years 3
Intelig	2016	Melanoma	Cancer	Baxter (Angels)	CTGov 2 ¹ EMA/EMA 2 ¹ EU/CTR 2 ¹	LTFU-CTP 1 LTFU-PA 1	Safety 1	5-9 years 4
Kymriah	2017	Acute lymphoblastic leukemia, diffuse large B-cell lymphoma, Multiple myeloma	Haematological cancers	Novartis	CTGov 2 ¹ EMA/EMA 1 ¹ EU/CTR 1 ¹	LTFU-CTP 3 LTFU-PA 1	Both 2	10-14 years 2
Libmeldy	2020	Neurochromic leukodystrophy	Rare diseases	Orion Therapeutics	CTGov 2 ¹ EMA/EMA 1 ¹ EU/CTR 1 ¹	LTFU-CTP 4 LTFU-PA 3	Both 1	10-14 years 3
Lucuma	2017	Leber's congenital amaurosis, retinitis pigmentosa	Rare diseases	Bayer Therapeutics (France)	CTGov 2 ¹ EMA/EMA 1 ¹ EU/CTR 2 ¹	LTFU-CTP 5 LTFU-PA 2	Efficacy 1	5-9 years 3
Rodanvec	2022	Hemophilia A	Haematology others	Baxter	CTGov 2 ¹ EMA/EMA 2 ¹ EU/CTR 2 ¹	LTFU-CTP 2 LTFU-PA 2	Both 1	10-14 years 1
Siborvelo	2016	Adenosine deaminase deficiency	Rare diseases	Orion Therapeutics	CTGov 2 ¹ EMA/EMA 1 ¹ EU/CTR 1 ¹	LTFU-CTP 4 LTFU-PA 1	Both 1	10-14 years 1
Tecartus	2020	Mantle cell lymphoma, acute lymphoblastic leukemia	Haematological cancers	Kia Pharma (Spain)	CTGov 2 ¹ EMA/EMA 1 ¹ EU/CTR 2 ¹	LTFU-CTP 4 LTFU-PA 1	Both 5	10-14 years 1
Uplesca	2022	Amnion Leukemia and decarboxylase deficiency	Rare Diseases	Pfizer Therapeutics	CTGov 2 ¹ EMA/EMA 1 ¹ EU/CTR 1 ¹	LTFU-CTP 1 LTFU-PA 1	Both 1	10-14 years 1
Yescarta	2017	Diffuse large B-cell lymphoma, non-Hodgkin's lymphoma, Multiple myeloma	Haematological cancers	Kia Pharma (Spain)	CTGov 2 ¹ EMA/EMA 1 ¹ EU/CTR 2 ¹	LTFU-CTP 11 LTFU-PA 2	Both 10	10-14 years 8
Zilgerna	2019	Spinal muscular atrophy	Rare diseases	Astell (Netherlands)	CTGov 2 ¹ EMA/EMA 1 ¹ EU/CTR 2 ¹	LTFU-CTP 5 LTFU-PA 2	Both 6	10-14 years 4

¹See QR code for additional information on these products.
Abbreviations: CTGov = ClinicalTrials.gov; CTP = clinical trial patients; EU-CTR = EU Clinical Trials Registry; HMA-EMA = Health of Medicines Agency-European Medicines Agency Catalogue of RWD Studies; LTFU = long-term follow-up; PA = post-authorization.

Methods

- Data on ongoing and completed LTFU-PA and LTFU-CTP studies with at least 5 years of follow-up was extracted from three sources: ClinicalTrials.gov (CTGov),¹ HMA-EMA Catalogue of RWD (HMA-EMA)² and EU Clinical Trials Register (EU-CTR)³ using search strings in **Supplemental Table 2** (see quick response (QR) code). Studies were excluded if one of the following applied:
 - Studies withdrawn or terminated
 - Studies included GCTs as add-on therapies, not as the investigational drug
 - Registration for multi-therapy disease registries
 - The survey aims to measure the effectiveness of additional risk minimization measures.
- After deduplication from the three sources, data was extracted from included studies on:
 - Objectives: safety, effectiveness, or both
 - Data sources: primary data collection (PDC), secondary data use (SDU), or hybrid
 - Presence of a comparator group with hypothesis testing: Yes or No
 - Target sample size
 - Length of follow-up, categorized as 5-9 years, 10-14 years, and 15+ years
 - Source population by care setting: post authorization (commercial setting), clinical trial (phase I to IV), or both
 - Whether pediatric populations were included
 - EMA's risk management plan (RMP) category (only for studies recorded in the HMA-EMA)
- Study characteristics were evaluated over time, and simple bivariate analyses were performed for study counts.

Figure 1. Long-term Follow-up Studies Identified from CTGov, HMA-EMA, and EU-CTR

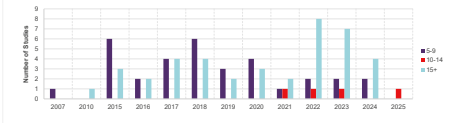


Abbreviations: CTGov = ClinicalTrials.gov; EU-CTR = EU Clinical Trials Registry; HMA-EMA = Health of Medicines Agency-European Medicines Agency Catalogue

Results

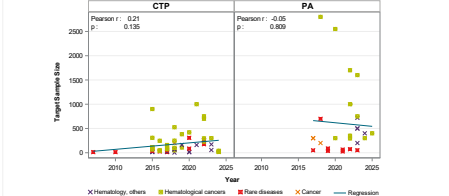
- Of 200 screened studies, 77 were selected for in-depth analysis (Figure 1), including three that evaluated more than one GCT. Most of these studies started in 2020 or later (n=39, 51%), and only two before 2021.
- Study NCT02016477 to evaluate Lucuma's long-term safety was the first one, starting in 2021.⁴
- Yescarta (n=13), Abema (n=10), and Caryfil (n=8) account for almost 40% of all studies, partly explaining the predominance of hematological cancers (n=38, 49%), followed by other hematological conditions (n=18, 23%), and rare diseases (n=17, 22%).
- A majority of studies evaluated both safety and effectiveness (n=66, 86%), focused on patients treated in clinical trials (n=50, 65%), followed patients for 15 years or longer (n=40, 52%), were based on primary data collection (n=6, 8%), did not include comparative inferential analyses (n=68, 88%), or did not include pediatric populations (n=50, 65%).
- All studies that follow that patients, and most studies following patients treated in both settings (n=2, 3%) use PDC, and almost half of studies following only patients treated in commercial settings leverage secondary data either exclusively (n=6, 25%) or in hybrid designs (n=5, 21%).
- All shorter studies (5-9 years) leverage PDC, but a majority of the longer ones (over 15 years) use existing registries, either exclusively (n=32, 49%) or in hybrid designs (n=4, 6%).
- All SDU studies are for hematological diseases, and 50% of those with hybrid designs are for rare diseases.
- All studies in rare diseases included pediatric patients, as did the majority of those based on hybrid data (n=4, 67%).
- None of the studies focusing only on safety and very few designed for both types of objectives (n=8, 12%) were designed for comparative analyses, but 33% of effectiveness studies were.
- 17 of the 18 studies identified in the HMA-EMA catalogue required a risk management plan (RMP). Category 1 RMPs, imposed as a condition of marketing authorization, accounted for the majority of studies (n=14, 82%), followed by category 2 (specific obligation of marketing authorization) and one requiring category 3. Most studies had safety and effectiveness objectives (n=11, 61%), and none aimed at evaluating only effectiveness. The vast majority planned to follow up with patients for 15 years or longer (n=13, 76%).

Figure 2. Distribution of Length of Follow-up by Study Start Year



- After 2020, significantly more studies followed up patients for 15 years or longer, and a new category aiming for a minimum 10-year follow up emerged.

Figure 3. Relationship between Study Initiation Year and Sample Size



- Figure 3 depicts starting year of LTFU-CTP (left panel) and LTFU-PA (right panel) studies, in relations with sample size.
- The first LTFU-PA study started in 2017. There are fewer of them, compared to LTFU-CTP studies (27 vs 63).
- LTFU-PA studies targeted larger samples and consistently over time, but the target sample size of LTFU-CTP studies increased slightly, albeit the trend was not statistically significant.
- LTFU studies in hematological cancers had larger sample sizes compared to studies in other therapeutic areas, and they could reach up to 1,000 patients for LTFU-CTP and 2,000 for LTFU-PA studies.

Conclusions

- 2020 seems to be an inflection point, after which some features seem to be converging, most notably the length of follow-up. This change is probably driven by the introduction of FDA's "Long-Term Follow-Up After Administration of Human Gene Therapy Products".⁵ While we could not distinguish studies designed for submissions to EMA or FDA, sponsors are likely gathering evidence for both from several of the studies assessed.
- In indications where robust clinical registries exist, data is used and only sometimes supplemented. Existing and de novo clinical registries are valuable sources for mandated LTFU studies.
- Fostering pre-competitive collaborations to create new or support upgrades of existing multi-sponsor long-term registries, could support consistent and sustainable research programs as part of an integrated LTFU evidence generation strategy, that includes patients treated with GCTs in both clinical trials or routine clinical practice.
- HMA-EMA catalogue and, even more so, the EU-CTR could benefit from better reporting in the structured fields. Facilitating the systematic revision of past and ongoing studies would enable a positive learning feedback for all stakeholders involved.
- The number of studies is not sufficient for a more in-depth analysis of most categories, but we uncovered trends that will hopefully inform the design of future studies.

References

1. European Medicines Agency. Council of the EU. 2007. Reg (EC) No 1394/2007 of the European Parliament and of the Council of 20 December 2007 on advanced therapy medicinal products and amending Directive 2001/83/EC of the European Parliament and of the Council on medicinal products for human use. <https://eur-lex.europa.eu/eli/reg/2007/1394/oj>
2. EMA. 2020. Guideline on quality, non-clinical, and clinical requirements for investigational ATMPs in clinical trials. https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-on-quality-non-clinical-and-clinical-requirements-for-investigational-atmps-in-clinical-trials_en.pdf
3. EMA. 2020. Guideline on safety and efficacy follow-up and risk management of ATMPs. https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-on-safety-and-efficacy-follow-up-and-risk-management-of-atmps_en.pdf
4. EMA. 2020. Guideline on quality, non-clinical, and clinical requirements for investigational ATMPs in clinical trials. https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-on-quality-non-clinical-and-clinical-requirements-for-investigational-atmps-in-clinical-trials_en.pdf
5. US FDA. 2020. "Long-Term Follow-Up After Administration of Human Gene Therapy Products". <https://www.fda.gov/oc/long-term-follow-up-after-administration-of-human-gene-therapy-products>
6. NCT02016477. <https://clinicaltrials.gov/ct2/show/study/NCT02016477>

Abbreviations: NAB and PR2 are employees of PPD's Evidence+Real-World Data & Scientific Solutions. Thermo Fisher Scientific is a registered trademark of Thermo Fisher Scientific. Thermo Fisher Scientific, Thermo Fisher Scientific logo, and PPD are trademarks of PPD Observational Studies, Thermo Fisher Scientific, Thermo Fisher Scientific logo, and PPD. All other trademarks are the property of their respective owners. © 2025 PPD. All rights reserved. This document is for informational purposes only and does not constitute an offer of any financial product or service. The information contained herein is not intended to be used as a basis for investment decisions. The information contained herein is not intended to be used as a basis for investment decisions. The information contained herein is not intended to be used as a basis for investment decisions.


Poster Title

Review of Ongoing PASS Submitted to the HMA-EMA Catalogue of RWD Studies Evaluating Cancer as an Outcome

Objective

- Summarize the key characteristics and methodologic approaches of post-authorization safety studies (PASS) evaluating cancer as an outcome that have been submitted to the Heads of Medicines Agencies (HMA)-European Medicines Agency (EMA) Catalogue of RWD Studies.

<p>AUTHORS</p> <p>Melissa Moore-Clingenpeel Leslie R Harrold Nicole Foster Heather J. Litman</p>	<p>AFFILIATION</p> <p>Thermo Fisher Scientific, Waltham, MA, USA</p>	<p>PRACTICE AREA</p> <p>PPD™ CorEvitas™ Clinical Registries, Thermo Fisher Scientific</p>	<p>CLIENT</p> <p>INTERNAL</p>
---	---	--	--------------------------------------



C-150

Review of Ongoing PASS Submitted to the HMA-EMA Catalogue of RWD Studies Evaluating Cancer as an Outcome

Melissa Moore-Clingenpeel,¹ Nicole Foster,¹ Leslie R Harrold,¹ Heather J. Litman¹
¹Thermo Fisher Scientific, Waltham, MA, USA

Background

- It is challenging to evaluate cancer as a safety outcome due to the long latency period for development and detection.
- Real-world data (RWD) are increasingly being leveraged to explore potential carcinogenic effects of medications used for the treatment of non-cancerous conditions. These data reflect real-world patterns of drug use and employ large, representative samples of patients receiving treatment as part of routine care over a prolonged follow-up period.
- In 2014, the FDA and National Cancer Institute (NCI) identified the following key themes for consideration to help address the inherent challenges with RWD:^{1,2}
 - Identifying cancer outcomes within existing data sources using computer algorithms
 - Evaluating all cancers in aggregate versus investigating specific cancers
 - Defining risk window and latency
 - Estimating risk
 - Identifying a comparison cohort
 - Evaluating rare exposures and rare cancers
- Ten years later, it is unclear how widespread and carefully these themes have been adopted in research studying cancer as an adverse event.

Objectives

- Summarize the key characteristics and methodologic approaches of post-authorization safety studies (PASS) evaluating cancer as an outcome that have been submitted to the Heads of Medicines Agencies (HMA)-European Medicines Agency (EMA) Catalogue of RWD Studies

Methods

HMA-EMA Catalogue of RWD Studies—Formerly PAS Register³

- The search criteria shown in Figure 1 were employed to identify candidate studies.
- Two independent reviewers collected data from each study on: year the study started, whether it was a regulatory requirement, exposure definition, follow-up duration, outcome definition, analytic approach, and other study characteristics.
- Only studies with submitted protocols were retained for analysis.

Results (cont.)

Study Design Elements

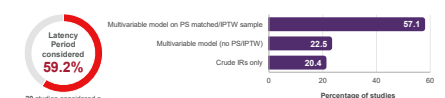
- Most studies (81.2%) followed patients for at least 5 years after initiation of a study-eligible therapy. Notably, however, more than 18% of studies did not mention in the PASS protocol how long patients would be followed (Figure 4).
- More than one-third of studies (38.8%) only planned to evaluate cancer in the aggregate versus studying specific types of cancers (e.g., lymphoma, lung cancer, breast cancer, etc.) (Figure 4).
- No studies were designed to specifically detect rare cancers.
- 16.3% of studies included prevalent use of the study treatment(s) (previously started and ongoing) at the study index date (Figure 4).

Figure 4. Follow-up duration, outcome definition, and handling of prevalent use



- More than half of the studies (59.2%) included an analysis with a latency period to account for delayed development and detection of cancer, typically 6 or 12 months. Most commonly, these were included as sensitivity analyses (Figure 5).
- When considering the analytic approach to evaluate cancer risk between study cohorts, 20.4% of the studies did not adjust for confounding or selection bias. Conversely, more than half of the studies employed propensity score (PS) matching or inverse probability of treatment weighting (IPTW) to balance cohorts at baseline, with extra adjustment in subsequent multivariable models (Figure 5).
- Except for studies reporting only crude incidence rates (IRs), nearly all studies used Cox regression; the remainder used Poisson regression.


Figure 5. Consideration of cancer latency and primary analytic approach



Abbreviations: IPTW = inverse probability of treatment weighting; IR = incidence rate; PS = propensity score

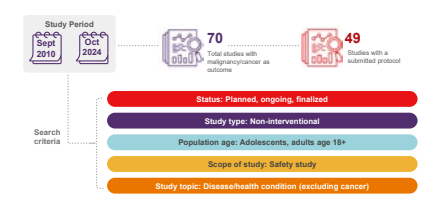
- Once-exposed always-at-risk exposure definitions (OEAR) were most common, accounting for 73% of all primary approaches (Figure 6, Table 1); subtypes included intention to treat (ITT), overlapping, and OEAR-drug of interest (OEAR-DOI).

Figure 6. Schematics for OEAR exposure definitions



Abbreviations: ITT = intention to treat; OEAR = once-exposed always-at-risk; OEAR-DOI = once-exposed always-at-risk drug of interest

Figure 1. Overview of search criteria and studies included



Search criteria:

- Status: Planned, ongoing, finalized
- Study type: Non-interventional
- Population age: Adolescents, adults age 18+
- Scope of study: Safety study
- Study topic: Disease/health condition (excluding cancer)

Figure 2. Characteristics of included studies




Figure 3. Populations of interest for included studies

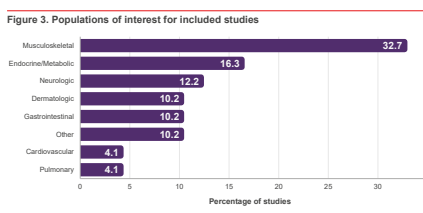


Table 1. Exposure definitions employed in included studies

Exposure Definition	Primary Definition N (%)	Any Mention ^a N (%)
Not noted	3 (6.1)	3 (6.1)
OEAR, unspecified type	6 (12.2)	7 (14.3)
ITT	10 (20.4)	12 (24.5)
Overlapping	10 (20.4)	11 (22.5)
OEAR-DOI	10 (20.4)	10 (20.4)
On treatment or per protocol	6 (12.2)	20 (40.8)
Durational/dose	1 (2.0)	11 (22.5)
Other exposure	3 (6.1)	5 (10.2)

Abbreviations: ITT = intention to treat; OEAR = once-exposed always-at-risk; OEAR-DOI = once-exposed always-at-risk drug of interest.
^aAny mention includes all mentions of the exposure definition across the primary analysis and all sensitivity, secondary, and/or exploratory analyses

Limitations

- This systematic review only included studies submitted to the HMA-EMA Catalogue of RWD Studies with a study protocol attached; it may not be representative of all studies evaluating cancer as an outcome.
- Some studies were designed prior to 2014 when the FDA/NCI recommendations were discussed; however, no notable differences in key design elements were seen in studies started before vs. after 2014.

Conclusions

- There is considerable heterogeneity in approaches to evaluating cancer as an outcome in PASS.
- Implementation of the FDA/NCI recommendations remains limited.


References

- US FDA. Public Meeting. September 10, 2014.
- Pharmasoft. *et al.* *Ann Epidemiol.* 2016;26(1):173S-40.
- HMA-EMA. Available at: <https://catalogues.ema.europa.eu>

Conflicts of Interest and Acknowledgments

Funding for this poster was provided by Thermo Fisher Scientific. PPD™ CorEvitas™ Clinical Registries have been supported through contractual relationships to the last two years by AbbVie, Amgen, Inc., Astra, Biologics Integrity, Bristol Myers Squibb, Cytex, Eli Lilly and Company, Genentech, Janssen Pharmaceutics, Inc., LEO Pharma, Novartis, Otsuka Pharmaceuticals, Pfizer, Inc., Sanofi Pharmaceuticals, Inc., and UCB. All authors are employees of PPD CorEvitas and did not receive outside funding to conduct the study. Editorial and graphic design support were provided by Caroline Cole, and Heather Nakagawa of Thermo Fisher Scientific.

Presented at the 41st International Society for Pharmacoepidemiology (ISPE) Annual Meeting | 22–26 August 2025 | Washington, DC, USA



Poster Title

Vaccine-related Studies Listed in the Heads of Medicines Agencies-European Medicines Agency (HMA-EMA) Catalogue of Real-world Data Studies: What Can We Learn from Study Characteristics?

Objective

- To provide an overview of key design characteristics of vaccine-related studies recorded in the Heads of Medicines Agencies-European Medicines Agency (HMA-EMA) Catalogue of RWD studies.

<p>AUTHORS</p> <p>Sarah Rosen¹ Ariel Berger³ Maria-Angeles Natividad-Sancho¹ Alice Rouleau¹ Sara Anglemann²</p>	<p>AFFILIATION</p> <p>¹Thermo Fisher Scientific, Lyon, France ²Thermo Fisher Scientific Stockholm, Sweden ³Thermo Fisher Scientific, Waltham, MA, USA</p>	<p>PRACTICE AREA</p> <p>PPD™ Observational Studies, Thermo Fisher Scientific</p>	<p>CLIENT</p> <p>INTERNAL Chalmers University and University of Gothenburg</p>
---	---	---	--

Vaccine-related Studies Listed in the Heads of Medicines Agencies-European Medicines Agency (HMA-EMA) Catalogue of Real-world Data Studies: What Can We Learn from Study Characteristics?



Poster Number

Sarah Rosen¹, Maria-Angeles Natividad-Sancho¹, Sara Anglemann², Ariel Berger³, Alice Rouleau¹
¹Thermo Fisher Scientific, Lyon, France; ²Thermo Fisher Scientific Stockholm, Sweden; ³Thermo Fisher Scientific, Waltham, MA, USA

Background

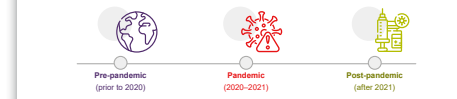
- The number of diseases that can potentially be prevented—or whose impacts could be mitigated—due to vaccination have increased over the past years.¹
- The number of commercially launched vaccines continues to increase—to cover “new” diseases as well as to provide options for coverage against existing diseases and their variants.
- While information on their efficacy/effectiveness and safety has always been of interest to decision-makers, the importance of assessing these attributes and the urgency of doing so in an expedient manner that nonetheless produces valid information have increased since the onset of the COVID-19 pandemic.
- Studies based on real-world data (RWD) are of particular importance and are highly relevant, as they illustrate the expected impact of vaccination in routine clinical practice (vs. the highly regulated environments of clinical trials) and represent the only way by which true effectiveness and safety of vaccines can be assessed at a population level and for the target population when compared with the limited sample size and generalizability of randomized controlled trials.

Objectives

- To provide an overview of key design characteristics of vaccine-related studies recorded in the Heads of Medicines Agencies-European Medicines Agency (HMA-EMA) Catalogue of RWD studies.

Methods

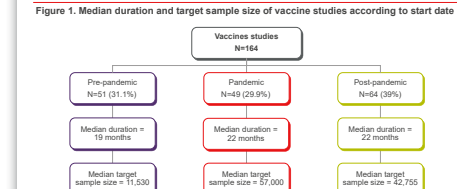
We conducted a search of all vaccine-related studies contained within the HMA-EMA Catalogue of RWD studies published from the beginning of database capture (2007) until 08 January 2025 and exported all such entries to identify and characterize their main design characteristics. The search term used for the search was “vaccine.” Each identified study was then categorized by vaccine type (using manual review as appropriate in instances where study entries (if the information was missing in the “product field”) to categorize studies into different vaccine types. Studies were stratified based on date of start of the study, as follows:



Manual review of study description or other variables was performed as relevant when inconsistencies or missing data was identified. For example, when filtering only on “safety” in the field related to study scope, the number of studies seemed low; we therefore investigated the classification options, and when adding all studies classified as risk minimization measure studies that have a safety objective, the number of studies with a safety outcome increased. Furthermore, when studies aimed to assess the occurrence of a particular safety event, these were also included in the overall total.

Results

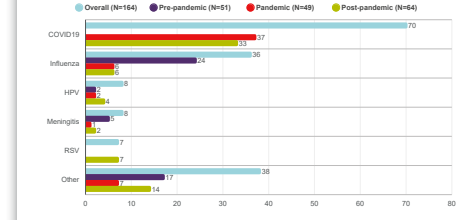
Median duration and target sample size of vaccine studies according to start date



Most (57.3%), or 84 of the 144 identified studies) studies were conducted as a requirement by a regulatory body, and 63 (38.4%) were EU Risk Management Plan (RMP) category 3 (required); 105 (84.0%) studies had private funding.

Types of vaccines studied and scope of studies

Figure 2. Type of vaccines studied according to start date

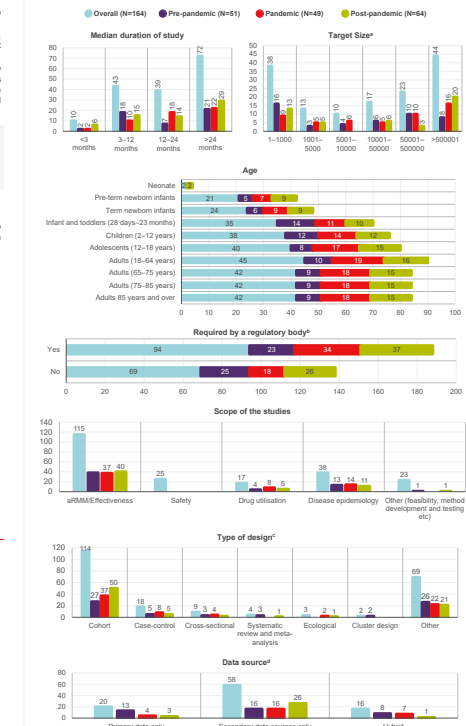


Abbreviations: HPV = human papillomavirus; RSV = respiratory syncytial virus. The catalogue also reported studies evaluating vaccines for pertussis (seven studies), malaria (six studies), herpes zoster (six studies), and rotavirus (four studies). Of note, nine studies did not specifically evaluate a particular vaccine but rather assessed feasibility of assessing background incidence or evaluation of data sources for the identification of global vaccine safety signals for example. The primary uses of real-world studies were to assess vaccine effectiveness (115 studies, 70.1%) and safety (55 studies, 38.4%) studies. The most common study designs were cohort studies (114, 69.5% studies) followed by case-control studies (18 studies, 11.0%). There was a trend toward increasing use of secondary data over time (43.2% secondary data only studies in 2007–2019 vs. 73.7% in 2020–2024).

Presented at the 41st International Society for Pharmacoepidemiology (ISPE) Annual Meeting | 22–26 August 2025 | Washington, DC, USA

Results (cont.)

Figure 3. Study design characteristics among 164 observational vaccine-related studies



*Nineteen studies did not report the overall sample size. †One study reported “unknown” for this variable. ‡Five missing for type of design. ††Eight missing for data source.

Limitations

- As noted previously, several data elements (e.g., study drug, medical condition to be studied) were missing or information were sparse, hindering our ability to search and characterize identified studies—the review of multiple fields of data within the catalogue was necessary to more accurately define study characteristics and demonstrates the limit of evaluating these only based on one criterion.
- Furthermore, we also identified inconsistencies in the data reported per study. To improve the quality of the catalogue, a review of the data prior to publication in the catalogue could add value and avoid errors in reporting.

Conclusions

- The post-marketing vaccine-related studies included in the HMA-EMA Catalogue employ various study designs and data collection approaches, with an increased focus on the value of RWD to address important questions related to vaccine effectiveness and safety.
- From the results, we can see that:
 - The use of these studies has increased greatly since the COVID-19 pandemic.
 - The willingness and appreciation of the ability of secondary data to address these key questions has grown over time, as has the number of designs employed.
 - The “RD” design is not explicitly mentioned in the list of designs—this may be classified under the case-control design or under other (the “other” category was highly used during reporting).
 - There is a call to action for better capture of study types/designs to help inform understanding and research on this topic.

References

- Marlow E. Vaccine preventable diseases are making “vaccines comeback.” *Health leaders warn.* *BMJ.* 2023;385:e073.
- © 2025 Thermo Fisher Scientific. All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, without the prior written permission of Thermo Fisher Scientific. This document is for informational purposes only and does not constitute an offer of any product or service. All trademarks are the property of their respective owners.

Presented at the 41st International Society for Pharmacoepidemiology (ISPE) Annual Meeting | 22–26 August 2025 | Washington, DC, USA



Poster Title

Methodological Considerations about Policy Impact on Studies of Pregnancy Registries and Maternal/Infant Outcomes

Objective

Objective 1

- Identify potential types of bias resulting from the Dobbs decision in the collection and/or analysis of pregnancy data in the US (Table 1 and Table 2).

Objective 2

- Determine which methods have been used to adjust for those sources of bias and assess their appropriateness (Table 3).

<p>AUTHORS</p> <p>Kathleen (Shannon) Torman Ronna L Chan Yao Yao Syd Phillips</p>	<p>AFFILIATION</p> <p>Thermo Fisher Scientific, Waltham, MA, USA</p>	<p>PRACTICE AREA</p> <p>PPD™ Observational Studies, Thermo Fisher Scientific</p>	<p>CLIENT</p> <p>INTERNAL</p>
--	---	---	--------------------------------------

Methodological Considerations about Policy Impact on Studies of Pregnancy Registries and Maternal/Infant Outcomes



B-206

Kathleen (Shannon) Torman¹, Yao Yao¹, Ronna L Chan¹, Syd Phillips¹
¹Thermo Fisher, Waltham, MA, USA

Background

- The Dobbs ruling, a US Supreme Court decision that the Constitution does not protect the right to induced abortion, may introduce bias to the analysis and interpretation of pregnancy registry data.
 - For example, the ruling may increase the proportion of subjects who are lost to follow-up and adversely affect registry enrollment and data completeness.
 - Statistical analyses of US pregnancy registry data in the post-Dobbs era should identify and adjust for sources of bias resulting from this decision.
- Pregnancy registry studies are designed to measure the effects of exposure to a study drug or vaccine on maternal and/or fetal pregnancy outcomes, including elective terminations, spontaneous abortions, and congenital malformations.
- These data points are usually obtained from provider data, self-reported from patients, or a combination of the two. The integrity of the analysis ultimately depends on the accuracy and completeness of the source data.
- Following the Dobbs decision, several states passed legislation to restrict their residents from accessing abortion services.
 - For example, in Texas, abortion is banned (with narrow exceptions) to preserve the life or a "major bodily function" of the mother, private citizens can sue abortion providers for services provided after six weeks of pregnancy,¹ and doctors can be punished with up to "99 years in prison, fines and the loss of their medical license" for terminating a pregnancy.²
 - The New York Times reported that 19 states passed legislation to ban abortion. Of those, 12 implemented a near total ban, and the remaining seven implemented policies with gestational limits ranging from 6 to 18 weeks of pregnancy.
- Therefore, it is important for investigators studying fetal outcomes to consider whether state-specific abortion policies may introduce one or more type(s) of bias to pregnancy registry data.

Objectives

- Objective 1**
Identify potential types of bias resulting from the Dobbs decision in the collection and/or analysis of pregnancy data in the US (Table 1 and Table 2).
- Objective 2**
Determine which methods have been used to adjust for those sources of bias and assess their appropriateness (Table 3).

Methods

- We performed a targeted literature review (TLR) in PubMed using the following terms related to the Dobbs decision: induced abortion, bias, methods, and pregnancy registries.

Results

- The initial TLR identified 192 potential articles (Figure 1). Filtering the search to articles published after the Dobbs decision reduced the number of articles to 37 (Figure 1). The results were further narrowed down to select articles that focused on identifying and/or adjusting for sources of bias in US pregnancy data.
- After the results were refined, four articles were confirmed to be truly relevant.³⁻⁶ These articles identified several potential sources of bias, including loss to follow-up, selection bias, reporting bias, and misclassification of abortion interventions and/or outcomes.
- Only one of the four articles conducted a statistical analysis to assess the impact of bias on reporting treatment effect. A covariate associated with the bias (loss to follow-up) and the outcome (treatment failure) was identified and then was used to conduct multiple imputations for missed treatment outcomes. This simulation-based method evaluated how the exclusion of those with unknown outcomes may have biased the estimated treatment effects.

Table 1. Targeted Literature Search for Objective 1

Step	Actions Taken	Number of Articles Identified
1	Searched PubMed using the MESH terms "induced abortion"[All Fields] AND ("bias"[MeSH Terms] OR "bias"[All Fields]) AND ("method"[All Fields] OR "methods"[MeSH Terms] OR "methods"[All Fields] OR "method"[All Fields] OR "methods"[MeSH Subheading])	192
2	Filtered the above results to articles within the past 5 years	37
3	Narrowed the results to select articles that focused on identifying and/or adjusting for sources of bias in US pregnancy data	4

Results (cont.)

Table 2. Types of Bias Identified (Objective 1)

Type of Bias	Associated Articles (See References)
Confounding	3
Selection Bias	3, 5
Misclassification	3
Protocol deviations	3
Missing data	3
Measurement of Outcomes Bias	3
Reporting Bias	3, 4, 5
Randomization Bias	3
Loss to Follow-up	6

Figure 1. Targeted Literature Search for Objective 1



Table 3. Targeted Literature Search for Objective 2

Search Actions Taken	Number of Articles Identified
Narrowed results from steps 1–3 identified in Table 1 and further restricted to articles that conducted a statistical analysis to	1

Conclusion

- The Dobbs ruling can introduce potential bias to pregnancy registry studies. Appropriate statistical methods should be considered to adjust for those biases in the analysis to correctly interpret the study outcomes.

References

- McCann A, Walker AS. Tracking Abortion Laws Across the Country. *New York Times*. July 2025. Accessed August 2025. <https://www.nytimes.com/interactive/2024/us/abortion-laws-roe-v-wade.html>
- Aldrige O. New Texas Law Aims to Save Lives by Clarifying the State Abortion Ban. Will it Work? *National Public Radio*. July 2025. <https://www.npr.org/2025/07/16/1616161616/1-5445143/bias-abortion-life-of-mother>
- Bernard M, Niemann J, Weisbold L, Jepsen D, Misai C, Luck-Sikorski C. Abortion stigma among abortion seekers, healthcare professionals and the public in high-income countries: A mixed-methods systematic review protocol. *BMJ open*. 2024;14(1):e078602.
- Cely-Aranda L, Cárdenas-García K, Enriquez-Santander LC, Saavedra-Avendano B, Avelaño GAO. Telemedicine for the provision of medication abortion to pregnant people at up to twelve weeks of pregnancy: a systematic literature review and meta-analysis. *Reproductive Health*. 2024;21(1):136.
- Mason H, Wolpin A, Goode B, McKenna C, Wilkinson T. Recruiting pregnant people considering abortion: An analysis of differences in participant characteristics by recruitment site across online, abortion fund, and clinic-based sources. *Contraception*. 2023;128:110139.
- Rosenberg JD, Raymond E, Shochet T. An examination of loss to follow-up and potential bias in outcome ascertainment in a study of direct-to-patient telemedicine abortion in the United States. *Contraception*. 2023;122:109996.

Disclosures

KB, RLC, and SP are employees of PPD™ Observational Studies, Thermo Fisher Scientific. YY was an employee of Thermo Fisher Scientific at the time that this poster was submitted. Funding for this poster was provided by Thermo Fisher Scientific.

Acknowledgements

Editorial and graphic design support were provided by Caroline Cole and Kaeleth Nakajima of Thermo Fisher Scientific.

Poster Title

An Alternative to Non-responder Imputation for Study Participants with Missing Outcomes in Single-arm Drug-effectiveness Studies: A Case Study in the PPD™ CorEvitas™ Psoriasis Registry

Objective

- To demonstrate an alternative to NRI for binary outcomes in a single-arm study that allows evaluation of more clinically informed assumptions about true response rates for patients with missing outcomes.

<p>AUTHORS</p> <p>Michael Singleton Oksana Pugach Wanyi Chen Robert R McLean</p>	<p>AFFILIATION</p> <p>Thermo Fisher Scientific, Waltham, MA, USA</p>	<p>PRACTICE AREA</p> <p>PPD™ CorEvitas™ Clinical Registries, Thermo Fisher Scientific</p>	<p>CLIENT</p> <p>INTERNAL</p>
---	---	--	--------------------------------------

An Alternative to Non-responder Imputation for Study Participants with Missing Outcomes in Single-arm Drug-effectiveness Studies: A Case Study in the PPD™ CorEvitas™ Psoriasis Registry



Michael Singleton, Oksana Pugach, Wanyi Chen, Robert R McLean, Nicole Middaugh, Heather J Litman, and Leslie R Harrold
Thermo Fisher Scientific, Waltham, MA, USA

Background

- Single-arm drug-effectiveness studies commonly assess patient responses to a drug of interest within a specific follow-up period. For example, the PPD™ CorEvitas™ Psoriasis (PsO) Registry collects patient and provider information during visits at 6-month intervals.¹
- A proportion of study-eligible patients will lack a qualifying follow-up visit. Excluding these patients from effectiveness analyses can introduce selection bias and undermine the validity of study findings.²
- Clinician experts tell us anecdotally that the most common reason for missing a 6-month follow-up visit is treatment success.
- Non-responder imputation (NRI) is a common approach for binary outcomes that imputes non-response among all patients with missing outcomes; however, it is not always a clinically sound assumption.
- Hedeker and colleagues proposed an alternative to NRI for missing smoking status at follow-up in smoking-cessation research, which can be adapted for single-arm pharmacoeconomics studies.³

Objectives

- To demonstrate an alternative to NRI for binary outcomes in a single-arm study that allows evaluation of more clinically informed assumptions about true response rates for patients with missing outcomes

Methods

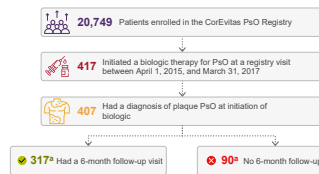
Source population

- The CorEvitas PsO Registry is a prospective, multicenter, non-interventional registry, launched in April 2015, for patients with PsO who are under the care of a dermatologist.
- The Registry collects data from both clinicians and patients at the time of routine outpatient clinical dermatological encounters.
- Registry inclusion criteria include age ≥ 18 years, written informed consent for participation, and start or switch to an eligible systemic PsO treatment at or within the previous 12 months of the date of enrollment.

Study population

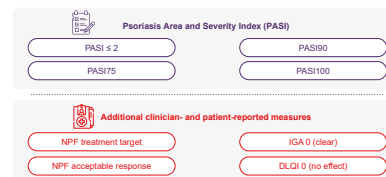
- Eligible patients for the current study were those diagnosed with plaque PsO who initiated a biologic therapy for PsO between April 15, 2015, and March 31, 2017, and were followed for a 6-month visit (0- to 9-month window) or were lacking a 6-month follow-up visit (Figure 1).

Figure 1. Patient selection flow diagram



* 56 of 317 patients with a 6-month follow-up visit and 24 of 90 patients with no 6-month follow-up visit discontinued index biologic before 6 months. Abbreviation: PsO = psoriasis.

Figure 2. Outcome measures



NPF treatment target: body surface area (BSA) ≤ 1%. NPF acceptable response: BSA ≤ 3% or 75% improvement from baseline. Abbreviations: DLQI = Dermatology Life Quality Index; IGA = Investigator Global Assessment; NPF = National Psoriasis Foundation.

Statistical analysis

- For patients without an outcome (Figure 2), the number of responders was singly imputed assuming (based on expert knowledge) that the most common reason for missing this visit is that the patient was responding well to therapy and chose not to return.
- This assumption was represented by odds ratios (ORs) of 1.5, representing a moderate increase in odds of responding to therapy, and 3.0, representing a large increase, relative to those with a 6-month follow-up visit.
- For patients who discontinued the index biologic before (or without) the 6-month follow-up visit, 6-month outcomes were carried forward from their most recent visit on or before the discontinuation date.
- Proportions of responders, and 95% CIs, were calculated first using complete cases only, then for all patients using imputation under each assumed OR. For N=66 subjects without a 6-month follow-up visit, who did not discontinue during the 9-month study period, the number who would achieve the outcome under the assumed OR was calculated.
- For comparison, results were also calculated under NRI.

Results

- The study population was a mean age of 50.1 years. Approximately half of the patients were women, and 83% were White. Differences between patients with and without a 6-month follow-up visit were observed in disease characteristics such as mean PASI, mean body surface area (BSA), mean Investigator Global Assessment (IGA), and mean Dermatology Life Quality Index (DLQI) (Table 1).
- The imputation methods resulted in a nearly 20% increase in sample size compared to the omission of patients without follow-up.

Conflicts of Interest and Acknowledgments

This study was funded by Thermo Fisher Scientific. PPD™ CorEvitas™ Clinical Registries have been supported through contracted subscriptions to the last 2 years by AbbVie, Amgen, Inc., Astra, Boehringer Ingelheim, Bristol Myers Squibb, Celis, Eli Lilly and Company, Genentech, GSK, Janssen Pharmaceuticals, Inc., Leo Pharma, Novartis, Ortho Dermatologics, and Thermo Fisher Scientific. PPD™ CorEvitas™ Clinical Registries are supported by the National Psoriasis Foundation (NPF). All authors are employees of PPD™ CorEvitas™ Clinical Registries, Thermo Fisher Scientific, and did not receive outside funding to conduct this study. O.P., M.L., and L.R. are also employees of Thermo Fisher Scientific. Medical writing support was provided by Kelly Shieh, PhD, and website and graphic design support was provided by Candice Cota, and Heather Nealegus of Thermo Fisher Scientific.

Results cont.

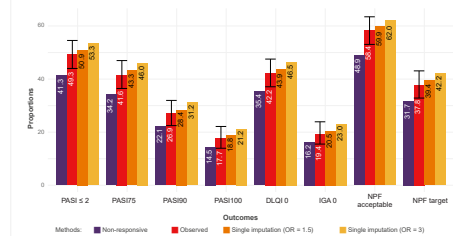
Table 1. Baseline characteristics of systemic biologic initiators

Characteristics	All Initiators N = 407	Initiators With 6-month Follow-up Visit n = 317	Initiators Without 6-month Follow-up Visit n = 90	SMD
Sociodemographic characteristics				
Age (years), mean (SD)	50.1 (14.2)	50.2 (14.3)	49.6 (14.0)	0.044
Female, n (%)	206 (50.6%)	158 (49.8%)	48 (53.3%)	0.070
Race, n (%)				0.147
White	338 (83.0%)	264 (83.3%)	74 (82.2%)	
Black/African American	18 (4.4%)	13 (4.1%)	5 (5.6%)	
Asian	30 (7.4%)	22 (6.9%)	8 (8.9%)	
Other	21 (5.2%)	18 (5.7%)	3 (3.3%)	
Disease characteristics				
Psoriatic arthritis, n (%)	185 (46.8%)	144 (47.1%)	41 (46.1%)	0.020
Duration of psoriasis disease (years), mean (SD)	16.2 (13.4)	16.2 (13.5)	16.4 (12.9)	0.016
PASI, mean (SD)	9.0 (7.9)	8.6 (7.5)	10.4 (8.1)	0.216
BSA, mean (SD)	15.3 (16.5)	14.8 (16.8)	16.8 (18.2)	0.114
IGA, mean (SD)	2.9 (0.9)	2.9 (0.9)	3.0 (0.8)	0.152
DLQI, mean (SD)	1.9 (1.1)	1.8 (1.2)	2.0 (1.1)	0.175

SMD compares initiators with and without a 6-month follow-up visit. Other race includes Native Hawaiian/Pacific Islander, American Indian/Alaska Native, Multiracial, Other. Abbreviations: BSA = body surface area; DLQI = Dermatology Life Quality Index; IGA = Investigator Global Assessment; PASI = Psoriasis Area and Severity Index; SMD = standardized mean difference.

- Observed patient responses varied across outcomes (Figure 3, Table 2A & 2B). Observed and imputed responses are shown in Figure 3.

Figure 3. Percentage of patients who achieved each outcome at 6-month follow-up under different methods for handling missing data



n = 334 for PASI ≤ 2, PASI75, and PASI100, and n = 341 for all other outcomes. These n's include 317 subjects with 6-month follow-up visit plus 24 without a 6-month follow-up visit who discontinued during the study follow-up period. These were 7 subjects for whom PASI responses could not be calculated due to missing or zero PASI at baseline. Abbreviations: DLQI = Dermatology Life Quality Index; IGA = Investigator Global Assessment; NPF = National Psoriasis Foundation; CR = odds ratio; PASI = Psoriasis Area and Severity Index; SMD = standardized mean difference.

Table 2A. Observed proportions achieving PASI response at 6-month follow-up

Outcome Measure (n = 334)	n (%)	95% CI
PASI75	139 (41.6%)	[0.36, 0.47]
PASI90	90 (26.9%)	[0.22, 0.32]
PASI100	58 (17.7%)	[0.14, 0.22]

Table 2B. Observed proportions of clinical outcome measures at baseline and 6-month follow-up

Outcome Measure (n = 341)	Baseline		6-month Follow-up	
	n (%)	95% CI	n (%)	95% CI
PASI ≤ 2	45 (13.2%)	[0.10, 0.17]	108 (30.3%)	[0.44, 0.56]
NPF treatment target	22 (6.5%)	[0.04, 0.10]	129 (37.8%)	[0.33, 0.43]
NPF acceptable response	57 (16.7%)	[0.13, 0.21]	199 (58.4%)	[0.53, 0.63]
IGA 0	12 (3.5%)	[0.02, 0.06]	66 (19.4%)	[0.16, 0.24]
DLQI 0	47 (13.8%)	[0.11, 0.19]	144 (42.2%)	[0.37, 0.48]

95% confidence interval (CI) calculated using Wilson score interval without continuity correction. NPF acceptable response at baseline defined as BSA ≤ 3%. DLQI 0 (no effect) at baseline was 345 due to 1 missing observation. Abbreviations: DLQI = Dermatology Life Quality Index; IGA = Investigator Global Assessment; NPF = National Psoriasis Foundation; PASI = Psoriasis Area and Severity Index.

Limitations

- A key limitation of this approach is the use of single imputation to estimate the total number of subjects who would achieve the outcome under the assumed OR. With single imputation, it is unclear how to properly estimate a CI around the point estimate.

Conclusions

- The findings illustrate how investigators can incorporate subject matter expertise and clinical experience into the specification of assumptions about the responses of patients with missing data on binary outcomes.
- Future work will extend the method to incorporate multiple imputations and different outcome types.

References

- Kramer JM. Clin Exp Rheumatol. 2016;34(5 Suppl 1):158-69.
- Hedeker D, et al. Epidemiology. 2016;27(1):9-17.
- Hedeker D, et al. Addiction. 2007;102:1564-73.



Poster Title

Landscape and Study Design Characteristics of Prospective Observational Cohort Studies on Pregnant Women in China

Objective

- To describe the landscape and key study design characteristics of prospective, observational, pregnancy cohort studies in China.

AUTHORS

Yufei Song^{1,2}
Alice Rouleau³
Syd Philips⁴

AFFILIATION

¹Thermo Fisher Scientific, Beijing, China
²Hainan Lecheng Institute of Real-World Study and Hainan University, Hainan, China
³Thermo Fisher Scientific, Paris France
⁴Thermo Fisher Scientific, Waltham, MA, USA

PRACTICE AREA

PPD™ Observational Studies,
Thermo Fisher Scientific

CLIENT

INTERNAL
Hainan Lecheng Institute of
Real-World Study and Hainan
University

Landscape and Study Design Characteristics of Prospective Observational Cohort Studies on Pregnant Women in China

ThermoFisher
SCIENTIFIC

B-266

Yufei Song^{1,2}, Alice Rouleau³, Syd Philips⁴

¹Thermo Fisher Scientific, Beijing, China; ²Hainan Lecheng Institute of Real-World Study and Hainan University, Hainan, China; ³Thermo Fisher Scientific, Paris France; ⁴Thermo Fisher Scientific, Waltham, MA, USA

Background

- Prospective observational cohort studies on pregnant women in China have evolved significantly in recent years, reflecting a growing recognition of the importance of maternal and child health. However, these studies are not frequently funded by industry.
- The evaluation of drug and medicinal products in pregnant women lacks guidelines from the Center of Drug Evaluation of the National Medical Products Administration, contrasting with the pregnancy registry study guidelines/requirements established by the US FDA and the EMA for post-approval safety assessments.

Objectives

- To describe the landscape and key study design characteristics of prospective, observational, pregnancy cohort studies in China.

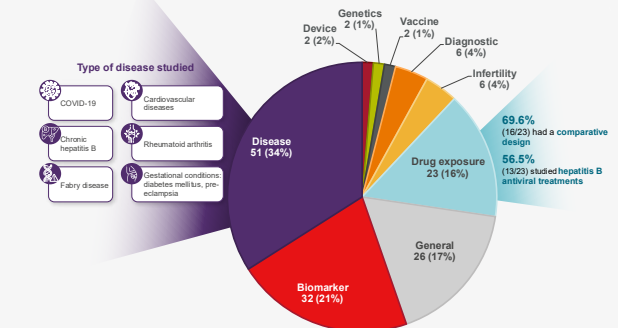
Methods

- We conducted a targeted search for prospective observational cohort studies of pregnant women in China in clinicaltrials.gov and in the Chinese Clinical Trial Registry (ChiCTR www.chictr.org.cn) as of 22 January 2025. Key study design characteristics were extracted through a manual review.

Results

- Of the 150 studies identified, there were 51 (34.0%) disease-specific, 32 (21.3%) biomarker studies, 26 (17.3%) general pregnant women cohorts, 23 (15.3%) drug-specific, and 18 other types, including diagnostic, genetics, vaccine, and device-specific studies (Figure 1).
- The indications vary across the 51 disease-specific studies, including pre-eclampsia, COVID-19, cardiovascular diseases, chronic hepatitis B, rheumatoid arthritis, Fabry disease, and gestational diabetes mellitus (Figure 1).

Figure 1. Type of Pregnancy Cohorts (n [%])



- The number of studies has increased since 2012, with 59.3% initiating between 2019 and 2024 (Figure 2).
- More than half (16/23, 69.6%) of the drug-specific studies were comparative. Thirteen (13/23, 56.5%) studies evaluated safety and effectiveness of antiviral treatments to prevent hepatitis B virus transmissions (Figure 1).

Figure 2. Study Start Year

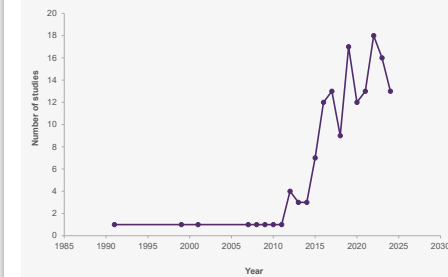


Figure 3. Cohort Sample Size Distribution in Log Scale



- The sample size varied across all types of studies with a median (Q1, Q3) of 900 (200, 3,768), whereas the median size of general pregnant women cohorts was higher, at 2,900 (1,620, 19,000) (Figure 3).
- The most common primary outcomes were maternal, fetal, and infant outcomes, including abortion, stillbirth, live birth, preterm birth, low birth weight, and birth defects.
- Almost all the studies were sponsored by government funding or self-funded by the affiliation of the principal investigator. Only six (4.0%) studies were funded by industry, including two disease cohorts (Fabry disease and Gaucher disease), two biomarker cohorts to predict pre-eclampsia and preterm birth, one hepatitis B virus treatment cohort, and one non-invasive device cohort for uterine fibroid treatment.

Disclosures

Y.S., A.R., and S.P. are current or former employees of PPD™ Observational Studies, Thermo Fisher Scientific. Funding provided by Thermo Fisher Scientific.

Acknowledgments

Editorial and graphic design support were provided by EDITOR'S NAME and Kawthar Nakayma of Thermo Fisher Scientific.

Conclusions

- The landscape of prospective observational pregnancy cohort studies in China has expanded significantly recently. However, industry involvement remains minimal, highlighting a need for clearer regulatory guidelines and increased industry engagement to enhance maternal and child health research specifically on medicinal products.

International Academy of Health Preference Research (IAHPR)

29 September–1 October 2025 | Enschede, Netherlands

Posters

Probabilistic Attribute Presentation in Discrete Choice Experiments: A Review of Current Practice

A Systematic Literature Review of Preference Studies in Migraine Treatments



Poster Title

Probabilistic Attribute Presentation in Discrete Choice Experiments: A Review of Current Practice

Objective

- This study aimed to synthesize current practice in presenting probabilities in DCEs, and report the nature of attributes, framing, and visual presentation.

AUTHORS

Gabriela S. Fernandez¹
Saudamini Oke²
Matthew Quaife^{2,3}

AFFILIATION

¹Thermo Fisher Scientific, Waltham, MA, USA
²London School of Hygiene and Tropical Medicine, London, UK
³Thermo Fisher Scientific, London, UK

PRACTICE AREA

PPD™ Evidera™ Patient-Centered Research, Thermo Fisher Scientific

CLIENT

INTERNAL
London School of Hygiene and Tropical Medicine



Probabilistic Attribute Presentation in Discrete Choice Experiments: A Review of Current Practice



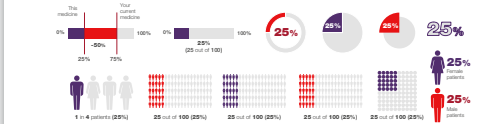
Gabriela S. Fernandez,¹ Saudamini Oke,² Matthew Quaife^{2,3}

¹Thermo Fisher Scientific, Waltham, MA, USA; ²London School of Hygiene and Tropical Medicine, London, UK; ³Thermo Fisher Scientific, London, UK

Background

- Discrete choice experiments (DCEs) are survey-based, quantitative preference elicitation tools used to measure trade-offs that people are willing to make between treatment or healthcare intervention attributes.
- Many DCEs need to represent probabilities to respondents; however, there are many potential ways to visually represent a probabilistic attribute (Figure 1). There is not consensus on how many probabilistic attributes can be considered simultaneously by participants.
- Risk communication research shows that individuals' understanding of risk can be highly variable.¹⁻⁷ FDA recommendations highlight the importance of ensuring that DCE attributes and levels are communicated such that they are understood consistently and effectively, such as avoiding fractions, decimals, different denominators, and relative scales, and encouraging the use of multiple presentation formats simultaneously.⁸
- Previous literature reviews have explored risk attribute presentation in DCEs⁹ and risk communication among patients with limited health literacy,¹⁰ finding that absolute attribute framing and use of visual aids in addition to verbal description improve comprehension.

Figure 1. Example Presentations for Presenting the Same Probability



Objectives

- This study aimed to synthesize current practice in presenting probabilities in DCEs, and report the nature of attributes, framing, and visual presentation.

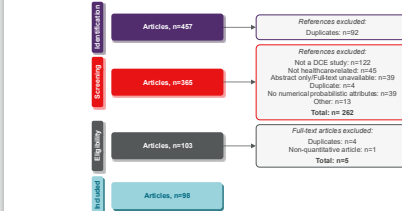
Methods

- Healthcare-related studies using DCEs were identified using databases including Medline, Embase, Web of Science, EconLit, and PsychINFO.
- Search results underwent title and abstract screening. Ten percent of abstracts and all full-text articles were double-screened by two independent researchers.
- Articles included in the review were stated preference studies published between October 2022 and August 2024, where DCEs asked participants to make a choice between two or more alternatives which included at least one probabilistic attribute.
- Data from included studies were extracted using a pre-specified framework.

Results

- Data from 98 studies were extracted (Figure 2).

Figure 2. PRISMA Flow Chart



- DCEs spanned a range of therapeutic areas, most commonly oncology (30%), endocrinology (11%), and dermatology (7%) (Table 1).
- Most studies included patients (83%), focused on assessing treatment preferences (81%), and reported some pre-testing (83%) (Table 1).
- The majority of DCEs included 5–7 attributes (70%) and 1–3 probabilistic attributes (77%) (Figure 3). Regardless of the total number of attributes, the average number of probabilistic attributes was around 2–3.

Figure 3. Number of Probabilistic Attributes by Total Number of Attributes



Results (cont.)

Table 1. Article Characteristics

Characteristic	N	%	Characteristic	N	%
Year of Publication			Therapeutic Area		
2022	8	8%	Oncology	29	30%
2023	59	60%	Endocrinology	11	11%
2024	31	32%	Dermatology	7	7%
Participant Type			Cardiology	6	6%
General Population	12	12%	Orthopedics	6	6%
Patients	81	83%	Gastroenterology	5	5%
Physicians	13	13%	Neurology	5	5%
Caregivers	4	4%	Respiratory	5	5%
Type of Healthcare Provision			Rheumatology	4	4%
Treatment	79	81%	Described Pre-testing/Piloting		
Disease Screening	4	4%	Reported	81	83%
Disease Prevention	10	10%	Not reported	17	17%
Disease Monitoring	3	3%			
Medical Technology	2	2%			

- More DCEs included probabilistic risk attributes (85%) than probabilistic benefit attributes (64%) (Table 2).
- Risks were most frequently presented as proportions alongside percentages (37%), and benefits most frequently as percentages only (30%) (Table 2).
- All risk attributes used absolute framing, whereas 16% of benefit attributes used relative framing (Table 2).
- Graphics were used more often to present risks than benefits (59% vs. 46%); within these, arrays of person-like figures were used in 72% of risks and 71% of benefits (Table 2).
- Exclusively male figures were used for 80% of benefits and 87% of risks. Blue (31%) and red (27%) colors were most used for risks, and blue (38%) and green (28%) colors were most used for benefits (Table 2).

Table 2. Summary of Presentation Across Risks and Benefits

Characteristic	Benefits*		Risks*		Characteristic	Benefits*		Risks*	
	N	%	N	%		N	%	N	%
Attribute Framing					Graphics Color*				
Absolute Framing	52	83%	83	100%	Blue	11	38%	15	31%
Relative Framing	10	16%	0	0%	Red	5	17%	13	27%
N/A	35	-	15	-	Black	2	7%	8	16%
Use of Graphic					Green	8	28%	2	4%
Yes	29	46%	49	59%	Yellow	0	0%	6	12%
No	28	44%	28	34%	Pink	2	7%	1	2%
Unknown	6	10%	6	7%	Orange	3	10%	5	10%
N/A	35	-	15	-	Purple	2	7%	5	10%
Graphic Type*					N/A	69	-	49	-
Icon Array	21	72%	42	86%	Numerical Presentation*				
Icon	4	14%	6	12%	Proportion with percentage	16	25%	31	37%
Circular Gauge	4	14%	2	4%	Bar Graph	2	7%	2	4%
Bar Graph	2	7%	2	4%	Proportion	8	13%	17	20%
N/A	69	-	49	-	Percentage	19	30%	10	12%
Icon Array Shape (if used)*					Proportion with percentage and description	4	6%	9	11%
Human (male)	12	57%	27	64%	Proportion with description	3	5%	8	10%
Human (ambiguous)	3	14%	3	7%	Percentage with description	15	24%	8	10%
Dots	3	14%	6	14%	N/A	35	-	15	-
Squares	2	10%	3	7%					
N/A	77	-	56	-					

Abbreviations: N/A = not applicable; *Not mutually exclusive; %n=83 studies included risks Pks and n=63 studies included benefits Pks.

Conclusions

- Current practices of presenting probabilistic attributes were generally consistent in following good practice in pre-testing and using some visual representation.
- The average number of probabilistic attributes was around 2–3 regardless of the total number of attributes; this may be the optimal amount of probabilistic information for participants to consider simultaneously.
- The variety of graphics used to display probabilistic risks and benefits, specifically gendershape, and color, indicate that work to understand how participants interpret different presentations may assist critical DCE design decisions.

References

- Kahneman D, Tversky A. *Cogn Psychol*. 1972;3(3):430-54.
- Birn J. *J Risk Uncertain*. 1997;14(3):301-6.
- Hanmatt JK, Graham JD. *J Risk Uncertain*. 1995;19(1):33-62.
- Lopes HA, et al. *Med Decis Making*. 2011;31(1):17-44.
- Ogden RG, et al. *Psychol Sci Public Interest*. 2007;8(2):53-96.
- Fischhoff B, et al. *Communications Risks and Benefits: An Evidence-based User's Guide*. U.S. Department of Health and Human Services, Food and Drug Administration; 2011.
- Gassman W, Ogden RG. *Z Evid Fortsch Qual Gesundheitsw*. 2008;10(7):111-3.
- Food and Drug Administration (FDA). *Incorporating Voluntary Patient Preference Information over the Total Product Life Cycle: Draft Guidance for Industry*. Food and Drug Administration Staff, and Other Interested Parties. 2024. <https://www.fda.gov/media/181000/download>
- Harrison M, et al. *Palmer*. 2014;7(2):151-70.
- Hocher R, et al. *Palmer Educ Course*. 2023;1(6):1078-84.

Disclosures

MO and GP are employees of PPD™ Evidera™ Patient-Centered Research, Thermo Fisher Scientific, a consultancy that provides scientific consulting services to pharmaceutical companies. The development of this poster was funded by Thermo Fisher Scientific.

Acknowledgments

Editorial and graphic design support were provided by Karen Treweek and Kavitha Nakayama of Thermo Fisher Scientific.



Poster Title

A Systematic Literature Review of Preference Studies in Migraine Treatments

Objective

- To identify, synthesize, and critically appraise stated-preference studies in migraine, with a focus on study design, treatment attributes and levels, analytical approaches, and the assessment of preference heterogeneity.

<p>AUTHORS Natasha Ramachandran¹ Harrison Clarke² Divya Mohan³ Jaein Seo²</p>	<p>AFFILIATION ¹Thermo Fisher Scientific, London, UK ²Thermo Fisher Scientific, Waltham, MA, USA ³Open Health, London, UK</p>	<p>PRACTICE AREA PPD™ Evidera™ Patient-Centered Research, Thermo Fisher Scientific</p>	<p>CLIENT INTERNAL Open Health, London, UK</p>
--	--	--	---

A Systematic Literature Review of Preference Studies in Migraine Treatments



Natasha Ramachandran,¹ Harrison Clarke,² Divya Mohan,³ Jaein Seo²
¹Thermo Fisher Scientific, London, UK; ²Thermo Fisher Scientific, Waltham, MA, USA; ³Open Health, London, UK

Background

- Migraines can be managed with acute, preventive, or combined treatments, yet care is often suboptimal due to undertreatment, variable clinical outcomes, and heterogeneity in the clinical course of migraine.
- Treatments differ in efficacy, tolerability, and administration, making individualized treatment plans that account for patient preferences essential.
- Incorporating patient preferences through shared decision-making can inform optimal treatment choices, enhance satisfaction, improve adherence, and guide the development of patient-centered treatments.
- A growing number of preference studies in migraine have sought to capture these insights, but vary in design, attribute selection, and analytical approach, limiting comparability and application. Despite this, no systematic synthesis exists.
- This systematic review fills that gap by consolidating evidence on patient preferences, identifying the treatment attributes most valued by relevant stakeholders, and highlighting evidence gaps, thereby providing a stronger foundation for patient-centered treatment strategies in migraine.

Objectives

- To identify, synthesize, and critically appraise stated-preference studies in migraine, with a focus on study design, treatment attributes and levels, analytical approaches, and the assessment of preference heterogeneity.

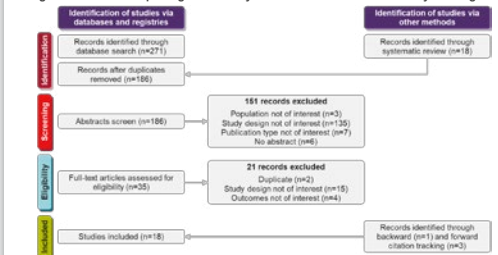
Methods

- Embase, MEDLINE, and the Cochrane Library were searched for relevant stated-preference studies on migraine treatments. Studies were eligible for review if they were original research articles employing stated-preference methods, including discrete choice experiment (DCE), best-worst scaling, thresholding, time trade-off (TTO), conjoint analysis, and contingent valuation method (CVM), to elicit preferences for treatments in migraine.
- Two researchers independently screened studies, and disagreements regarding the inclusion or exclusion of studies were resolved through discussion between the two researchers; any remaining discrepancies were adjudicated by a third researcher.
- Extracted information included study characteristics, methods for attribute and instrument development, choice task design, attribute framing, and analytical approaches.
- Narrative synthesis and descriptive statistics were used to summarize findings.

Results

- Two-hundred and seventy-one studies were identified from the literature search and screened. Eighteen studies were deemed eligible and included in the review (Figure 1).

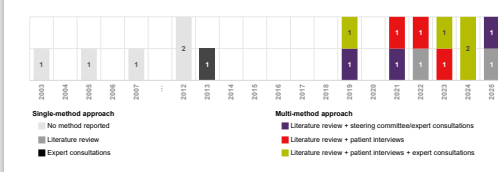
Figure 1. Preferred Reporting Items for Systematic Reviews and Meta-analyses Diagram



Backward citation tracking was conducted to identify articles cited by the review papers and forward citation tracking was conducted by looking at papers that were cited by the review papers to identify any missing publications.

- Ten studies employed multi-method approaches to develop preference elicitation instruments (Figure 2), including literature review combined with multiple steering committee or expert consultations (n=3),^{1,3} literature review combined with qualitative patient interviews or focus groups (n=3),⁴ and literature review combined with patient interviews and multiple expert and/or paper consultations (n=3).^{5,6}
- Three studies employed a single method (Figure 2): Literature review only (n=2)^{7,12} or clinical expert consultation only (n=1).¹³

Figure 2. Methods Used for Attribute and Instrument Development

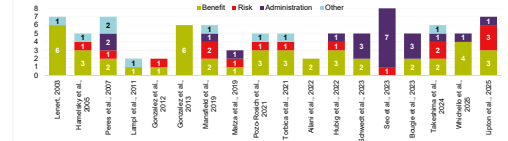


Disclosures

MR, NC, and JS are employees of PPD™ Evidera™ Patient-Centered Research, Thermo Fisher Scientific, a consultancy that provides scientific consulting services to pharmaceutical companies. The development of this paper was funded by Thermo Fisher Scientific.

Results (cont.)

Figure 3. Number of Attributes by Category across Included Studies



Benefit attributes

- Benefit attributes were the most frequently assessed, with 17 studies including at least one benefit concept and only one study omitting them entirely (Figure 3).
- The most commonly reported benefit concept was speed of onset (n=9), followed by durability of effectiveness (n=6), reduction in migraine/symptom severity (n=6), improvement in functional abilities (n=7), reduction in migraine/symptom frequency (n=7), consistency of effectiveness (n=3), and reduced need for rescue medication (n=1). Attributes were framed differently within the same concepts. For example, reduction in migraine or symptom severity was conceptualized as the presence or absence of specific symptoms,¹³ "nside/unchanged,"⁴ percentage reduction from the baseline³ or from the maximum,¹⁰ or percentage reduction of the overall migraine-specific symptoms.⁶
- Duration of effectiveness was expressed as how quickly treatment wears off,^{7,9} how likely¹³ or how quickly⁶ the headache returns, or how long symptom relief lasts.⁸
- Variation at the attribute level contributed to different framing approaches. For example, speed of onset levels were described as a percentage chance (e.g., 0%–22% chance of onset on the first day),¹¹ as time intervals in minutes and hours,^{6,14} or as days, weeks, or months.^{1,7,12}

Risk attributes

- Risk attributes were evaluated in 11 studies (Figure 3).
- The most commonly included adverse event (AE) was injection site reactions (n=5).
- Other risks included gastrointestinal effects, such as constipation (n=4),^{1,2,10,12} nausea (n=3),^{1,4,12} and stomach pain/discomfort (n=2),¹⁰ and cognitive effects (n=3),^{2,5,8} such as difficulty with memory and concentration.^{2,5,8}
- Four studies specified probabilities for distinct adverse events; others described risk in terms of reversibility (transient vs. permanent; n=1) or severity (n=1).

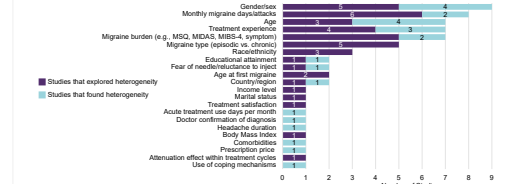
Administration and other attributes

- Mode and/or frequency of administration were the most common administration concepts, included in 10 studies (Figure 3). Modes of administration included oral pills (n=5), intravenous infusion (n=3),^{4,7,9} intramuscular injection (n=1),² subcutaneous injection (n=6),^{1,2,4,5,8,12} or unspecified biologics (n=1).⁴
- Other attributes included in the studies were cost (n=6), specifically out-of-pocket cost (Figure 3).

Preference heterogeneity

- Fourteen studies examined observable preference heterogeneity by sociodemographic and clinical characteristics, of which 11 identified characteristics associated with preference differences (Figure 4), including gender/sex (n=4),^{5,8,9,13} age (n=4),^{7,8,13} treatment experience (n=3),^{5,13} monthly migraine days (n=2),^{5,10} migraine burden (n=2),^{4,10,13} educational attainment (n=1),¹³ fear of needles (n=1),¹² country/region (n=1),⁴ acute treatment use days per month (n=1),¹⁰ physician confirmation of diagnosis (n=1),¹² headache duration (n=1),⁴ comorbidities (n=1),⁹ prescription price (n=1),⁸ and use of coping mechanisms (n=1).¹⁴

Figure 4. Respondent Characteristics Assessed for Preference Heterogeneity



Abbreviations: MBS-4 = Four-Item Migraine Burden Scale; MDAS = Migraine Disability Assessment; MSQ = Migraine-Specific Quality of Life

Conclusions

- This review highlights substantial variability in how treatment attributes were selected, framed, and analyzed in stated-preference studies in migraine. Greater methodological transparency, including using qualitative methods and patient-facing tools to improve rigor and reproducibility, is needed.
- Enhancing the assessment of preference heterogeneity and addressing cognitive burdens during instrument development are critical to improving the validity and real-world applicability.
- By supporting more robust study design and reporting, this review can help guide the conduct of higher-quality preference studies that better inform treatment decisions and development, ultimately improving migraine care.

References

- Lyons RB, et al. *Headache*. 2020.
- Matta LB, et al. *Qual Life Res*. 2015;24(9):2399-2372.
- Poon-Ross P, et al. *Pharm Ther*. 2020;20(10):1779-1793.
- Hugh LT, et al. *Headache*. 2022;62(5):1187-1197.
- Seo J, et al. *Health Econ*. 2023;12(2):181-198.
- Boyle A, et al. *J Environ Res Public Health*. 2021;18(9):4611-4621.
- Boyle A, et al. *Headache*. 2023;63(1):113-123.
- Martiniello C, et al. *Headache*. 2019;59(5):715-726.
- Schwartz T, et al. *Headache*. 2023;63(4):484-493.
- Swanson T, et al. *Pharm Ther*. 2024;13(1):161-163.
- Allen J, et al. *Headache*. 2022;62(3):374-382.
- Swanson T, et al. *Pharm Ther*. 2023;13(1):161-163.
- Gonzalez JM, et al. *Headache*. 2013;53(10):1635-1650.
- Hugh LT, et al. *Headache*. 2023;63(1):113-123.
- Hugh LT, et al. *Headache*. 2023;63(1):113-123.



GetReal

1–2 October 2025 | Utrecht, The Netherlands

Posters

Low-left truncation mitigation rates in post marketing pregnancy exposure registries assessing the risk of spontaneous abortion (SAB)

Integrating Real-world Evidence for Advanced Therapies in Inflammatory Bowel Disease into Reimbursement Submissions: Insights and Considerations from the UK

Use of Real-world Evidence in Advice Reports

Assessing Reimbursement in The Netherlands from 2023–2025

Is There a Lack of Dietary Data Collection in Real-world IBD Studies? Literature Review and Future Considerations

Are Cross-sectional Cohorts an Efficient Alternative to Prospective Cohort Design in Real-world Studies?

Data Reliability in Retrospective Chart Review Studies: Results and Considerations from a Novel Data Review Methodology

Cost-effectiveness of Radiotherapy in Uterine Serous Carcinoma (USC): A Real-world Study

A Real-world-derived Algorithm to Improve the Cost-effectiveness of External Beam Radiotherapy in Uterine Serous Carcinoma

Comparing Randomized Clinical Trials to Real-world Studies Evaluating the Effectiveness of a bDMARD in the Management of Crohn's Disease

Comparison of Randomized Controlled Trials (RCT) with Open-label, Single-cohort, Real-world Studies (RWS) in the evaluation of pharmacologic treatment of Non-small Cell Lung Cancer (NSCLC)

Utility of Real-world External Control Arm (ECA) in the Evaluation of Pharmacologic Treatment of Non-small Cell Lung Cancer (NSCLC)



GetReal 2025

1–2 October 2025 | Utrecht, The Netherlands

Poster Title

Low-left truncation mitigation rates in post-marketing pregnancy exposure registries assessing the risk of spontaneous abortion (SAB)

Objective

- To review the literature on approaches to the analysis of SAB risk in prospective observational studies, including pregnancy exposure registries.

<p>AUTHORS Maria Angeles Natividad Sancho¹ Reem Masarwa²</p>	<p>AFFILIATION ¹Thermo Fisher Scientific, Paris, France ²Thermo Fisher Scientific, Montreal, Canada</p>	<p>PRACTICE AREA PPD™ Observational Studies, Thermo Fisher Scientific</p>	<p>CLIENT INTERNAL</p>
---	--	--	--

Low-left truncation mitigation rates in post-marketing pregnancy exposure registries assessing the risk of spontaneous abortion (SAB)



ABS-01

Maria Angeles Natividad Sancho,¹ Reem Masarwa²
¹Thermo Fisher Scientific, Paris, France; ²Thermo Fisher Scientific, Montreal, Canada

Background

The risk for spontaneous abortion (SAB) in post-marketing safety studies is often evaluated through pregnancy exposure registries, where recruitment occurs after pregnancy recognition. However, most SABs occur early in pregnancy, often before pregnant women enroll into the registry,¹ and the risk for SAB rapidly decreases across the first trimester of pregnancy (left truncation). If not accounted for during the analyses, left truncation can lead to inaccurate conclusions or flawed decision-making.²⁻⁴

Objective

To review the literature on approaches to the analysis of SAB risk in prospective observational studies, including pregnancy exposure registries.

Methods

A structured review of the literature was conducted in July 2025 in the Heads of Medicines Agencies (HMA)/European Medicines Agency (EMA) Catalogue for real-world data (RWD) sources, Embase, and MEDLINE. The search was based on explicit keywords and eligibility criteria and used a search strategy that was developed in consultation with a research librarian. The goal was to summarize the methodology used to estimate the incidence and prevalence of SABs, as well as author-reported limitations related to the potential impact on current SAB epidemiologic estimates.

Peer-reviewed, full-text original research articles indexed in MEDLINE and Embase were included if they met the following criteria: (1) human study; (2) study designs: prospective; (3) reporting occurrence measures (prevalence, incidence, proportions); (4) reporting early pregnancy loss, spontaneous abortion, or miscarriage; (5) English-language studies; and (6) published between 1974 and July 2025 (Embase) and 1946 to July 2025 (MEDLINE). Commentaries, clinical trials, case series, case reports, and studies reporting data infertility, in vitro fertilization, or assisted reproductive techniques were excluded.

We also reviewed data from protocols and reports from all pregnancy registry studies registered in the HMA-EMA RWD Catalogues until July 2025 that were required by regulators.

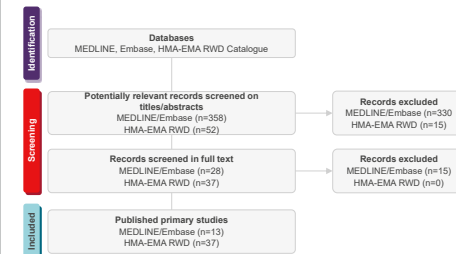
Titles, abstracts, and full texts of the potentially relevant articles and protocols/reports were screened by a single researcher to identify those that met the eligibility criteria.

Data on study characteristics, methodology, population, and outcomes of interest were extracted from eligible publications and protocols/reports into a bespoke spreadsheet. Limitations reported by the authors were extracted as free text and grouped by topic for the purposes of evaluation and discussion.

Results

Our search generated 52 potentially relevant protocols from pregnancy exposure registries from the HMA-EMA Catalogue for RWD sources and 358 peer-reviewed articles from MEDLINE/Embase. After screening, we included 37 (71%) pregnancy exposure registries published between 2011 and 2024, and 13 (4%) peer-reviewed manuscripts published between 2009 and 2023 (Figure 1).

Figure 1: Study search and screening flow diagram



Abbreviations: EMA = European Medicines Agency; HMA = Heads of Medicines Agencies; RWD = real-world data

Of the 37 pregnancy studies registered in the HMA-EMA Catalogue, most (78%) were required by regulatory authorities. When examining drug exposure, 35% reported SABs among pregnant women exposed to treatment for autoimmune diseases, 22% to vaccines, and 14% to antimigraine drugs. Sixty-five percent of the registries were conducted in North America and 27% in Europe; cohort sample sizes ranged from 50 to >1000 pregnancies.

Conclusions

- Left truncation, common to studies evaluating SAB rates, needs to be properly handled using survival analysis methods to avoid bias. Otherwise, when used in comparative analyses, the relative risk of SAB calculated from cohorts differentially impacted by this bias may be misleading.
- Only a small proportion of pregnancy registries required regulatory authorities, and an even smaller proportion of published studies that assessed SAB in the general population, used survival analysis methods to minimize left truncation bias when estimating SAB rates.
- A significant proportion of pregnancy registries, and fewer article reports, that did not account for left truncation in the analysis acknowledged it as a study limitation, suggesting there is a potential barrier to the use of survival analysis.

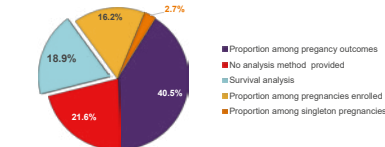
Presented at the GetReal Conference 2025 | 1–2 October 2025 | Utrecht, The Netherlands

Results (cont.)

HMA-EMA Catalogue pregnancy registries

Among the registries that described a SAB analysis (n=29), only 24% (n=7) implemented survival analysis methods to calculate SAB rates in one sample or 32 sample comparisons (e.g. exposed vs unexposed cohorts) (Figure 2). Of these, 57% (n=4) explicitly mentioned that survival analysis methods were used to handle left truncation, as well as right-censoring when a participant was lost-to-follow-up prior to 20 weeks' gestation or when the participant was no longer at risk of the event. An additional study (14%) performed a sensitivity analysis by gestational age at enrollment.

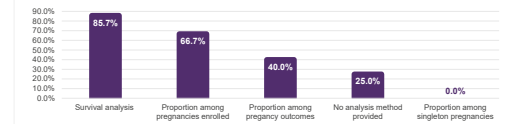
Figure 2: Calculation of SAB rates among HMA-EMA Catalogue pregnancy registries



Abbreviations: EMA = European Medicines Agency; HMA = Heads of Medicines Agencies; RWD = real-world data; SAB = spontaneous abortion

Most studies that used survival analysis to account for left truncation to estimate SAB incidence acknowledged the bias associated with any other type of analysis. Although to a lesser extent, bias was also acknowledged among studies that did not use survival analysis; overall, half of the reports that did not account for left truncation in the analysis acknowledged it as a study limitation (Figure 3).

Figure 3: Proportion of HMA-EMA Catalogue pregnancy registries that acknowledged bias

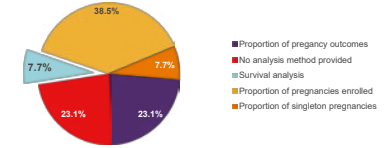


Abbreviations: EMA = European Medicines Agency; HMA = Heads of Medicines Agencies; RWD = real-world data

Embase/MEDLINE peer-reviewed articles on pregnancy studies

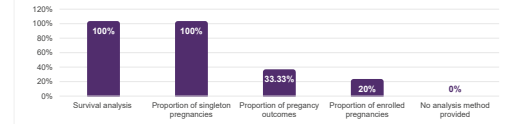
The source data for the peer-reviewed manuscripts included national pregnancy surveillance programs (n=3), pregnancy registries (n=3), national registry databases (n=3), and hospital-based cohorts (n=4). Only 8% of reports accounted for left truncation in the analysis (Figure 4), and 25% of those that did not account for left truncation by survival analysis acknowledged a potential methodological weakness (Figure 5).

Figure 4: Calculation of SAB rates among Embase/MEDLINE peer-reviewed articles on pregnancy studies



Abbreviations: EMA = European Medicines Agency; HMA = Heads of Medicines Agencies; RWD = real-world data; SAB = spontaneous abortion

Figure 5: Proportion of Embase/MEDLINE peer-reviewed articles on pregnancy studies that acknowledged bias—by analysis method type used



Abbreviations: EMA = European Medicines Agency; HMA = Heads of Medicines Agencies; RWD = real-world data

Disclosures
 MANS and RM are employees of PPD™ Observational Studies, Thermo Fisher Scientific. Funding for this poster was provided by Thermo Fisher Scientific.

Acknowledgements
 Editorial and graphic design support were provided by Caroline Cole and Kaefer Nakajima of Thermo Fisher Scientific.

© 2025 Thermo Fisher Scientific Inc. All rights reserved. All trademarks are the property of Thermo Fisher Scientific and its subsidiaries unless otherwise specified. This information is not intended to encourage use of these products in any manner that might violate the intellectual property rights of others.



Poster Title

Integrating Real-world Evidence for Advanced Therapies in Inflammatory Bowel Disease into Reimbursement Submissions: Insights and Considerations from the UK

Objective

- The objective was to evaluate RWE use in previous NICE HTA submissions for advanced IBD therapies for adults, including biologics and Janus kinase (JAK) inhibitors, and offer recommendations for incorporating real-world evidence (RWE) into future HTA submissions.

<p>AUTHORS</p> <p>Neil R. Brett¹ Garthiga Manickam¹ Roy Mootoosamy² Daniela Castano³ John S. Sampalis¹</p>	<p>AFFILIATION</p> <p>¹Thermo Fisher Scientific, Montreal, Canada ²Thermo Fisher Scientific, London, UK ³Thermo Fisher Scientific, Waltham, MA, USA</p>	<p>PRACTICE AREA</p> <p>PPD™ Observational Studies, Thermo Fisher Scientific PPD™ Evidera™ Health Economics & Market Access, Thermo Fisher Scientific</p>	<p>CLIENT</p> <p>INTERNAL</p>
--	---	--	--------------------------------------



Integrating Real-world Evidence for Advanced Therapies in Inflammatory Bowel Disease into Reimbursement Submissions: Insights and Considerations from the UK

Neil R. Brett,¹ Garthiga Manickam,¹ Roy Mootoosamy,² Daniela Castano,³ John S. Sampalis¹
¹Thermo Fisher Scientific, Montreal, Canada; ²Thermo Fisher Scientific, London, UK; ³Thermo Fisher Scientific, Waltham, MA, USA

ABS-19

Background

- Recent studies in inflammatory bowel disease (IBD) highlight notable differences between clinical trial populations and real-world study populations, including age, comorbidities, disease severity, and treatment responses.^{1,2} This may lead to limited generalizability of clinical trial results in real-world populations.
- Based on the flexibility of real-world study eligibility criteria, real-world data (RWD) may be able to provide insights into the patient experiences and outcomes of a broader population, which is crucial for making informed decisions about the efficacy and cost-effectiveness of treatments.
- Thus, the UK's National Institute for Health and Care Excellence (NICE) has a strategy to leverage RWD to bridge the knowledge gaps that exist in reimbursement submissions.³ However, it is not known if RWD was utilized in previous UK NICE health technology assessment (HTA) submissions for advanced IBD therapies.

Objective

- The objective was to evaluate RWE use in previous NICE HTA submissions for advanced IBD therapies for adults, including biologics and Janus kinase (JAK) inhibitors, and offer recommendations for incorporating real-world evidence (RWE) into future HTA submissions.

Methods

- A targeted review of NICE appraisals was performed for Crohn's disease (CD) and ulcerative colitis (UC) to understand RWE expectations, integration and usage.
- Available therapies for CD and UC were identified, focusing on those approved for adults. Appraisals for the approved therapies were retrieved from the NICE database. Each appraisal was independently reviewed by two individuals to extract information (year, indication, RWE utilization/context, and RWE study design.)

Results

- As of June 2025, eight biologics and three JAK inhibitors have been approved for adult CD and UC in Europe.

Figure 1. Approved therapies

	CD	UC
Biologics approved in adults	<ul style="list-style-type: none"> • Adalimumab • Infliximab • Risankizumab • Ustekinumab 	<ul style="list-style-type: none"> • Adalimumab • Mirikizumab • Vedolizumab • Ustekinumab • Guselkumab • Infliximab • Mirikizumab • Vedolizumab
JAK inhibitors approved in adults	<ul style="list-style-type: none"> • Upadacitinib 	<ul style="list-style-type: none"> • Filgotinib • Tofacitinib • Upadacitinib

Abbreviations: CD = Crohn's disease; IBD = inflammatory bowel diseases; JAK = Janus kinase; UC = ulcerative colitis

Figure 2. Our considerations for integrating RWE into HTA submissions (initial or updated after initial reimbursement)

- 1

Pre Product Approval

RWE could provide insights into patient populations for future trial planning:

 - To understand patient subgroups that may benefit most
 - Inform eligibility criteria that better reflect real-world clinical practice
 - Using patient voice to understand outcomes that matter most to patients
- 2

Pre through Post Approval

Using NICE guidelines to support RWE planning/conduct:

 - The NICE RWE framework guidance supports study design, including transparency, data quality and suitability, analytical rigor, and reproducibility
 - For treatment outcomes, guidelines help to define meaningful outcomes
- 3

Post Product Approval

Opportunities for patient-centered data collection for early insights:

 - Identify early data on treatment adherence and patient satisfaction with treatment
 - Early signals of adverse events and treatment tolerability
 - Understand time to response in real-world settings
- 4

Post Product Approval and Reimbursement

RWE could support treatment positioning after reimbursement:

 - Optimizing treatment sequencing for where a therapy may be most effective
 - Highlighting effectiveness in specific patient subgroups
 - Understanding treatment patterns to support cost-effectiveness conversations
 - Comparative effectiveness can influence clinical guidelines

Using NICE appraisals to support RWE planning/conduct:

 - To identify areas of clinical/cost-effectiveness uncertainty to guide RWE studies
 - To guide eligibility criteria for subgroups of interest

Abbreviations: NICE = National Institute for Health and Care Excellence; QoL = quality of life; RWE = real-world evidence

Conclusions

- RWE is under-utilized in HTA submissions to NICE for advanced therapies for IBD.
- Future submissions should consider knowledge gaps that could most benefit from RWE integration, including modeling assumptions.

References

1. Ingrassia Y, et al. *Pharmacol Res*. 2024 Feb;200:107074.
2. Vieujean S, et al. *Lancet Healthy Longev*. 2022 May;3(5):e356-e366.
3. NICE. NICE real-world evidence framework. 2022. Accessed October 19, 2023; <https://www.nice.org.uk/corporate/realworld/overview>.

Disclosures

NEB and JSS are employees of PPD™ Observational Studies, Thermo Fisher Scientific. GM and BM are employees of PPD Evidence™ Health Economics and Market Access, Thermo Fisher Scientific. DC is an employee of PPD Clinical research business of Thermo Fisher Scientific. Funding for this poster was provided by Thermo Fisher Scientific.

Acknowledgments

Editorial and graphic design support were provided by Michael Grossi and Shari Berger of Thermo Fisher Scientific.

Presented at the GetReal Conference 2025 | 1–2 October 2025 | Utrecht, The Netherlands

© 2025 Thermo Fisher Scientific, Inc. All rights reserved. All trademarks are the property of Thermo Fisher Scientific and its subsidiaries unless otherwise specified. This information is not intended to encourage use of these products in any manner that might violate the intellectual property rights of others.

GetReal 2025

1–2 October 2025 | Utrecht, The Netherlands

Poster Title

Use of Real-world Evidence in Advice Reports Assessing Reimbursement in The Netherlands from 2023–2025

Objective

- This study aimed to assess the use of RWE in submissions that are reviewed by ZIN for reimbursement advice.
- The goal was to create an overview and look at trends of RWE use and reimbursement recommendations.

AUTHORS

Janneke Luijken¹
Ivana Sestak²

AFFILIATION


¹Thermo Fisher Scientific, The Randstad, Netherlands
²Thermo Fisher Scientific, Milan, Italy

PRACTICE AREA

PPD™ Observational Studies,
Thermo Fisher Scientific
PPD Evidera™ Real-World Data
& Scientific Solutions, Thermo
Fisher Scientific

CLIENT

INTERNAL



ABS-20

Use of Real-world Evidence in Advice Reports Assessing Reimbursement in The Netherlands from 2023–2025

Janneke Luijken¹, Ivana Sestak²
¹Thermo Fisher Scientific, The Randstad, Netherlands; ²Thermo Fisher Scientific, Milan, Italy

Background

- In 2025, the EMA released a reflection paper on the generation of real-world evidence (RWE).
- The reflection paper describes how RWE can add to existing evidence from interventional studies, providing data on effectiveness and safety that is representative of clinical practice.¹
- The guideline published by Zorginstituut Nederland (ZIN) describes that RWE from observational studies can be included in economic evaluations and contribute larger external validity than interventional studies.²

Objectives

- This study aimed to assess the use of RWE in submissions that are reviewed by ZIN for reimbursement advice.
- The goal was to create an overview and look at trends of RWE use and reimbursement recommendations.

Methods

- A targeted review of advice reports published between January 2023 and July 2025 that evaluated new treatments for reimbursement in the basic insurance package was conducted in the ZIN database.³
- The reports included the submission with evidence provided by the market authorization holder (MAH).
- In the Netherlands, the basic insurance package is mandatory for all adults and covers all essential medical care.
- Relevant data was extracted from the submissions in the advice reports, including reimbursement recommendation decisions, type of RWE studies, and reasons for rejection. This data was summarized into a standardized data extraction form.
- The data are summarized with descriptive statistics, by presenting the number and proportion of advice reports in different categories.

Results

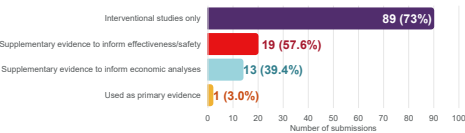
- A total of 122 advice reports were identified from the ZIN database.
- Figure 1 presents a high-level overview of the results of this targeted review.
- Of the 33 submissions that included RWE, 14 (42.4%) included data from retrospective studies, eight (24.2%) data from prospective studies, and six (18.2%) used real-world data as external control arms; for the remaining five submissions, type of RWE was not specified or data from databases/electronic health records were used.

Results (cont.)

Type of Evidence

- Most submissions included only data from interventional studies (N=89, 73.0%).
- The remaining 33 submissions (27.0%) included some type of RWE (Figure 3).
- Only one (3.0%) submission used RWE as the primary evidence.
- Nineteen (57.6%) submissions used RWE as supplementary evidence to inform effectiveness/safety.
- Thirteen (39.4%) submissions used RWE as supplementary evidence to inform economic analyses.

Figure 3. Type of Evidence in all Submissions, n (%)



Therapeutic Areas in Relation to RWE Use

- Half (50%) of the oncology submissions incorporated RWE in the submissions.
- In contrast, only 8.3% of neurology submissions included RWE.
- The five therapeutic areas with the highest percentage of RWE use in submissions is shown in Table 1.

Table 1. Top 5 Therapeutic Areas with RWE Use in Submissions

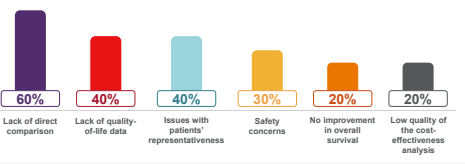
Therapeutic area (submissions)	Percentage of submissions including RWE
Oncology (n=38)	50.0%
Nephrology (n=2)	50.0%
Cardiology (n=5)	40.0%
Rare disease (n=24)	25.0%
Hematology (n=4)	25.0%

Abbreviation: RWE = real-world evidence

Reasons for Withholding Reimbursement Advice in Submissions with RWE

- The reasons for withholding reimbursement advice were summarized. Out of the 10 submissions that included some type of RWE but did not receive a reimbursement recommendation, "lack of a direct comparison" was listed in six advice reports.
- Figure 4 presents the different reasons listed in the 10 advice reports.

Figure 4. Reasons for Rejection of Reimbursement Advice



Conclusions

- Use of RWE in submissions in the Netherlands is relatively low, with 27% of submissions including RWE.
- Irrespective of type of data included (interventional or RWE), studies that include a comparator and quality of life assessments, with an unbiased representation to the general target population, are needed to support the evidence portfolio of new treatments.

References

1. EMA. Reflection paper on use of real-world data in non-interventional studies to generate real-world evidence for regulatory purposes. 2025. Accessed September 16, 2025. https://www.ema.europa.eu/en/medicines/regulatory-affairs/regulatory-affairs-articles/reflection-paper-on-use-of-real-world-data-in-non-interventional-studies-to-generate-real-world-evidence-regulatory-purposes_en.pdf
2. ZIN. Richtlijn voor het uitvoeren van economische evaluaties in de gezondheidszorg 2024. Accessed September 16, 2025. <https://www.zorginstituutnederland.nl/publicaties/publicatie/2024/07/18/richtlijn-voor-het-uitvoeren-van-economische-evaluaties-in-de-gezondheidszorg/Richtlijn-voor-het-uitvoeren-van-economische-evaluaties-in-de-gezondheidszorg-13284646-20240718.pdf>
3. ZIN. Publicatie. 2025. Accessed September 16, 2025. <https://www.zorginstituutnederland.nl/publicaties/publicatie/2025/07/16/publicatie-20250716.pdf>

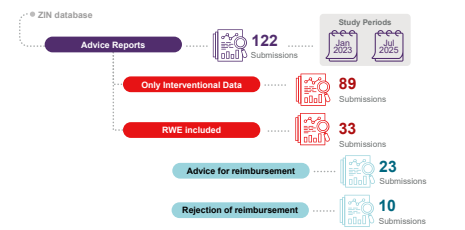
Disclosures

All authors are employees of PPD™ Observational Studies, Thermo Fisher Scientific. Funding provided by Thermo Fisher Scientific.

Acknowledgements

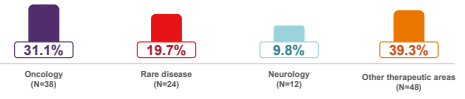
Editorial and graphic design support was provided by Michael Gross, Shari Berger, and Katherine Maloney of Thermo Fisher Scientific.

Figure 1. High-level Overview of Targeted Review Results




Abbreviations: RWE = real-world evidence, ZIN = Zorginstituut Nederland

Figure 2. Therapeutic Areas Assessed in Submissions



Presented at the GetReal Conference 2025 | 1–2 October 2025 | Utrecht, The Netherlands

© 2025 Thermo Fisher Scientific, Inc. All rights reserved. All trademarks are the property of Thermo Fisher Scientific and its subsidiaries unless otherwise specified. This information is not intended to encourage use of these products in any manner that might infringe the intellectual property rights of others.





GetReal 2025

1–2 October 2025 | Utrecht, The Netherlands

Poster Title

Is There a Lack of Dietary Data Collection in Real-world IBD Studies? Literature Review and Future Considerations

Objective

- The objective was to describe dietary data collection methods and purposes in observational IBD studies and provide considerations to enhance future research.

AUTHORS Neil R. Brett ¹ Garthiga Manickam ¹ Marielle Bassel ¹ Daniela Castano ² John S. Sampalis ¹	AFFILIATION ¹ Thermo Fisher Scientific, Montreal, Canada ² Thermo Fisher Scientific, Waltham, MA, USA	PRACTICE AREA PPD™ Observational Studies, Thermo Fisher Scientific PPD™ Evidera™ Health Economics & Market Access, Thermo Fisher Scientific PPD™ clinical research business of Thermo Fisher Scientific	CLIENT INTERNAL
---	--	---	---------------------------



Is There a Lack of Dietary Data Collection in Real-world IBD Studies? Literature Review and Future Considerations

Thermo Fisher Scientific

ABS-21

Neil R. Brett,¹ Garthiga Manickam,¹ Marielle Bassel,¹ Daniela Castano,² John S. Sampalis¹
¹Thermo Fisher Scientific, Montreal, Canada; ²Thermo Fisher Scientific, Waltham, MA, USA

Background

- Leading gastroenterology associations (e.g., American Gastroenterological Association [AGA] and the European Crohn's and Colitis Organization [ECCO]), have recent guidelines that highlight evidence showing that healthy dietary patterns in patients with inflammatory bowel disease (IBD) can help manage their disease, reduce progression, or induce remission.^{1,2}
- However, due to cost, data availability, and collection burden, it is unclear how often diet is ascertained in observational IBD studies.
- This gap assessment will help guide recommendations for integrating dietary data into real-world evidence and applying to clinical practice.

Objective

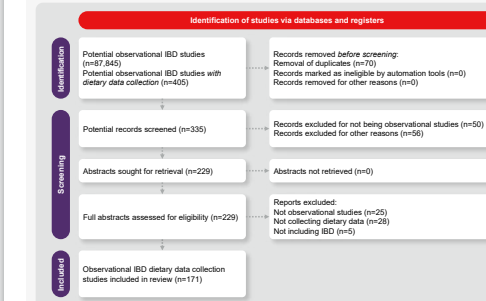
- The objective was to describe dietary data collection methods and purposes in observational IBD studies and provide considerations to enhance future research.

Methods

- A targeted literature review was conducted to identify observational IBD studies that collected dietary data. Searches were performed for IBD or ulcerative colitis (UC) or Crohn's disease (CD) and nutrition or diet or dietary or food or diet record or diet questionnaire.
- Records were retrieved from the databases and de-duplicated. Two individuals independently reviewed the remaining abstracts and full-text studies to ensure they met the search criteria and extracted key information, such as population, disease, data collection methods, and study outcomes.

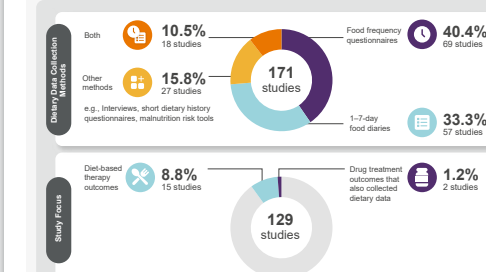
Results

Figure 1. PRISMA Flow Diagram of Abstract Screening



Abbreviations: IBD = inflammatory bowel disease; PRISMA = Preferred Reporting Items for Systematic Reviews and Meta-analyses.
• Of the 171 studies included, 40 (23.4%) studies were in CD, 24 (14.0%) were in UC, and 107 (62.6%) included both. One-hundred and nineteen studies (69.6%) focused on adults, while the remainder included children or both age groups.

Figure 2. Data Collection Methods

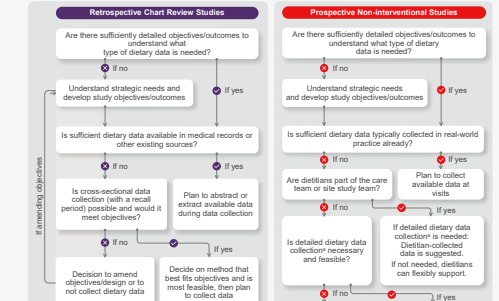


Key Finding (Figure 2): Despite known links between diet and IBD outcomes, only 1.2% of reviewed studies collected dietary data to support assessment of drug treatment effectiveness.

Conclusions

- Dietary data collection methods varied among observational IBD studies.
- Few studies collected dietary data when evaluating drug treatment outcomes.
- Future study designs should align data collection methods with research objectives/population and balance participant burden with data quality.
- Many real-world limitations can be minimized through planning—including early input from clinical, nutrition, and operational experts during study design and start-up.

Figure 3. Flow Diagrams For Selecting Dietary Data Collection Methods



Abbreviations: FFQ = Food Frequency Questionnaire. *Detailed dietary data collection could include 24-hour recalls or food diaries.

Figure 4. Considerations For Selecting Dietary Data Collection Methods

Alignment of Dietary Data Collection to Objectives/Outcomes	• To assess general dietary patterns (e.g., diet frequency, dietary quality) and association with outcomes • Use brief tools such as FFQs or short questionnaires	• To associate specific nutrients with disease outcomes • Use detailed methods (e.g., food diaries, multiple 24-hour recalls)	
Leveraging Existing Data Sources (When Possible)	• Patients may have some dietary data in clinical records: • Inpatients or those receiving enteral/parenteral nutrition • Pediatric or adult patients on structured dietary interventions • Patients with diagnosed malnutrition or under care of dietitians • Feasibility checks needed to determine level of detail available		
Cross-sectional Data Collection in Retrospective Studies	• Use validated FFQs: • Low burden, capture usual intake over the past 30 days to 1 year • Can be administered electronically or on paper	• Limitations: • Potential recall bias • FFQ recall period may not align with entire follow-up period • Requires patient consent	
Prospective Data Collection	• FFQs or brief dietary surveys at multiple timepoints: • Capture general intake patterns	• Food diaries (3–7 days) or 24-hour recalls: • Best for tracking detailed macronutrient intake • Require more participant effort	• Involving dietitians is strongly recommended: • Improves data quality • Supports feasible recall methods (e.g., person/phone) • Reduces participant burden
Technology Considerations	• Nutrition-specialized digital platforms: • Guided input (e.g., prompts, pictures) • Enables real-time data analysis and monitoring	• Emerging AI tools (e.g., food image recognition): • Improve data quality through visual verification • Allow proxy input (e.g., caregivers, parents) • May reduce patient burden	

Abbreviations: AI = artificial intelligence; FFQ = Food Frequency Questionnaire

References

- Hershah JG, et al. *Gastroenterology*. 2024 Mar;166(3):521-532.
- Vais Svolos V, et al. *J Crohns Colitis*. Published online July 12, 2025.

Disclosures

NRB and NDM are employees of PPD™ Observational Studies, Thermo Fisher Scientific. GM is an employee of PPD Evidera™ Health Economics & Market Access. DC is employed by PPD clinical research business of Thermo Fisher Scientific. Funding provided by Thermo Fisher Scientific.
Editorial and graphic design support were provided by Michael Grossi and Richard Lesson of Thermo Fisher Scientific.



GetReal 2025

1–2 October 2025 | Utrecht, The Netherlands

Poster Title

Are Cross-sectional Cohorts an Efficient Alternative to Prospective Cohort Design in Real-world Studies?

Objective

- The objective was to compare CSC study design with LPC study design in a cohort of patients with rheumatoid arthritis (RA).

<p>AUTHORS Neil R. Brett¹ Marielle Bassel¹ Vincent McCarty² John Sampalis^{1,2}</p>	<p>AFFILIATION ¹Thermo Fisher Scientific, Montreal, Canada ²McGill University, Montreal, Canada</p>	<p>PRACTICE AREA PPD™ Observational Studies, Thermo Fisher Scientific</p>	<p>CLIENT INTERNAL McGill University, Montreal</p>
---	--	--	---



Are Cross-sectional Cohorts an Efficient Alternative to Prospective Cohort Design in Real-world Studies?

Neil R. Brett,¹ Marielle Bassel,¹ Vincent McCarty,² John Sampalis^{1,2}

¹Thermo Fisher Scientific, Montreal, Canada; ²McGill University, Montreal, Canada

Background

- Longitudinal prospective cohort (LPC) study design may be considered the gold standard for real-world studies of change over time in clinical or patient-reported outcomes.
- However, these designs are sometimes non-feasible or cost prohibitive due to long study durations, patient burden, regulatory requirements, and/or logistical/operational complexities.^{1,2}
- Cross-sectional cohort (CSC) design, in which subjects are assessed once but at varying timepoints of their disease journey, may mitigate many of the design challenges posed by LPC studies.
- However, it is unknown how well CSC studies can replicate the outcomes demonstrated in LPC studies.
- Thus, research is needed comparing CSC with LPC study designs.

Objective

- The objective was to compare CSC study design with LPC study design in a cohort of patients with rheumatoid arthritis (RA).

Methods

- The study utilized real-world data from 1,716 patients with RA receiving either of two treatments.
- Clinical Disease Activity Index (CDAI) was assessed at 0, 3, 6, 9, and 12 months in the LPC study.
- To simulate the CSC study, a single timepoint was randomly selected as the hypothetical cross-sectional time of assessment.
- Univariate and repeated measures general linear models were used to analyze the data.
- Between-treatment differences were adjusted in both models for age, gender, duration of disease, prior treatments, and markers of inflammation.

Results

Table 1. Baseline characteristics

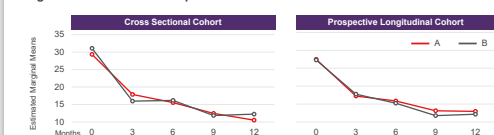
	Treatment		
	A	B	C
Age: Mean (SD) years	59.62 (11.46)	60.29 (11.61)	59.94 (11.54)
Gender N(%)			
Female	6,255 (73.90%)	617 (74.40%)	1272 (83.10%)
Male	231 (2.81%)	213 (25.7%)	444 (25.9%)
Disease Duration Mean (SD) years	7.16 (8.76)	8.02 (9.15)	7.59 (8.97)
CRP Mean (SD) mg/L	2.10 (3.55)	2.91 (3.05)	2.95 (3.32)
ESR Mean (SD) mm/hr	22.38 (17.13)	23.32 (16.63)	22.84 (16.89)
RF+ N(%)	484 (54.6%)	494 (59.5%)	978 (57.0%)
ACCP+ N(%)	284 (32.1%)	280 (33.7%)	564 (32.9%)
Prior Treatments			
Corticosteroids N(%)	211 (23.8%)	185 (22.3%)	396 (23.1%)
NSAIDs N(%)	268 (30.2%)	230 (27.7%)	496 (29.0%)
HQO N(%)	354 (40.2%)	338 (40.5%)	595 (40.2%)
Sulfasalazine N(%)	142 (16.0%)	140 (16.9%)	282 (16.4%)
Leflunomide N(%)	112 (12.8%)	116 (14.0%)	228 (13.3%)
Gold N(%)	7 (0.8%)	5 (0.6%)	12 (0.7%)
Biologic DMARD N(%)	359 (40.5%)	331 (39.8%)	690 (40.2%)

Abbreviations: ACCP+ = anti-cyclic citrullinated peptide-positive; CRP = C-reactive protein; DMARD = disease-modifying anti-rheumatic drug; ESR = erythrocyte sedimentation rate; HQO = hydroxychloroquine; NSAID = nonsteroidal anti-inflammatory drug; RF+ = rheumatoid factor-positive

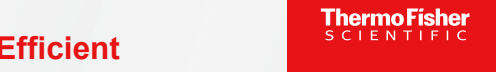
Table 2. Estimated least squared mean CDAI

Treatment	Month	Cross Sectional Cohort			Longitudinal Prospective Cohort				
		Mean	SE	95% CI	Mean	SE	95% CI		
A	0	29.38	0.971	27.459	31.270	27.49	1.975	23.619	31.365
	3	17.88	0.747	16.412	19.343	17.23	0.205	13.237	21.221
	6	15.55	0.729	14.125	16.983	15.86	2.118	11.704	20.012
	9	12.45	0.670	11.311	13.719	13.13	1.700	9.759	16.459
	12	10.52	1.003	8.553	12.488	12.95	1.861	9.297	16.597
B	0	31.10	0.997	29.140	33.051	27.31	1.976	23.437	31.190
	3	15.97	0.730	14.535	17.398	17.76	2.037	13.763	21.754
	6	16.17	0.732	14.735	17.606	15.24	2.120	11.098	19.401
	9	11.85	0.751	10.377	13.324	11.77	1.702	8.435	15.111
	12	12.24	1.039	10.207	14.281	12.15	1.863	8.494	15.800
Total	0	30.23	0.696	28.865	31.595	29.79	0.245	29.309	30.270
	3	16.92	0.522	15.898	17.947	16.20	0.253	15.702	16.895
	6	15.86	0.517	14.848	16.876	15.04	0.263	14.520	15.550
	9	12.15	0.503	11.161	13.135	12.09	0.211	11.679	12.505
	12	11.38	0.722	9.966	12.799	12.06	0.230	11.605	12.509

Figure 1. Estimated least squared mean CDAI



Presented at the GetReal Conference 2025 | 1–2 October 2025 | Utrecht, The Netherlands



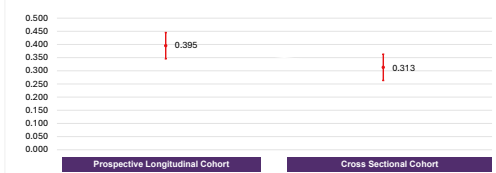
ABS-22

Table 3. Covariates

Variable	P-value	
	Cross-sectional Cohort	Longitudinal Prospective Cohort
Treatment	0.559	0.806
Month	0.001	0.001
Treatment * Month	0.136	0.655
Age	0.356	0.501
Gender	0.006	0.025
Disease Duration	0.051	0.025
Baseline CRP	0.061	0.003
Baseline ESR	0.764	0.734
RF	0.000	0.000
ACCP	0.055	0.001
Prior Treatments		
MTX	0.000	0.000
Corticosteroids	0.591	0.477
NSAIDs	0.001	0.000
HQO	0.008	0.079
Sulfasalazine	0.451	0.040
Leflunomide	0.935	0.161
Gold	0.281	0.957
Biologic	0.000	0.000

Abbreviations: ACCP+ = anti-cyclic citrullinated peptide; CRP = C-reactive protein; ESR = erythrocyte sedimentation rate; HQO = hydroxychloroquine; MTX = methotrexate; NSAID = nonsteroidal anti-inflammatory drug; RF = rheumatoid factor

Figure 3. Between-group differences



Discussion

- The results of this study showed that the CSC analysis (using general linear models) emulated the results of an LPC analysis (using a repeated measures mixed effects model).
- Between-group differences with the two analyses were similar.
- The two methods showed comparable statistical significance for most covariates included in the models.

Limitations

- CSC was simulated from a longitudinal study by selecting random visit numbers at different disease intervals.
- Because timing of assessment in real-world studies may be due to multiple known and unknown factors (i.e., may not be random), assessments of differential attrition and visit patterns would be required to ensure that the CSC model can be used. Adjustments for informed censoring may be necessary.
- This study was based on RA, a progressive chronic condition, where disease severity increases with time. Thus, the results may not apply to diseases where severity does not follow a linear trend.
- While the example was derived from RA, it is possible that the results could be generalized to other paradigms where longitudinal assessments of disease are needed.

Conclusions

- Cross-sectional analyses could be a valid replacement for longitudinal prospective analyses in some instances.
- A CSC approach offer advantages over the longitudinal studies with respect to duration and costs.
- Careful considerations must be given to potential issues of bias and suitability of the therapeutic area for utilization of the CSC approach.

References

- Aiurkawa D, et al. *JBMIR Plus*. 2018 Apr; 3(2):187-194.
- Saag KG. *Clinical Research Methods in the Rheumatic Diseases*. In: Firestein GS, Budd RC, Gabriel SE, et al, eds. *Kelly's Textbook of Rheumatology*, 9th ed. Elsevier; 2013:495-510.

Disclosures

MRB, MB and JB are employees of PPD™ Observational Studies, Thermo Fisher Scientific. Funding for this poster was provided by Thermo Fisher Scientific.

Acknowledgments

Editorial and graphic design support were provided by Michael Gossau and Shari Berger of Thermo Fisher Scientific.

© 2025 Thermo Fisher Scientific Inc. All rights reserved. All trademarks are the property of Thermo Fisher Scientific and its subsidiaries unless otherwise specified. This information is not intended to encourage use of these products in any manner that might infringe the intellectual property rights of others.



Poster Title

Data Reliability in Retrospective Chart Review Studies: Results and Considerations from a Novel Data Review Methodology

Objective

- The objective was to describe an alternative method for data quality review and highlight key considerations.

AUTHORS Neil R. Brett ¹ Luke Safarczyk ² Elizabeth Donahue ² Marielle Bassel ¹	AFFILIATION ¹ Thermo Fisher Scientific, Montreal, Canada ² Thermo Fisher Scientific, Waltham, MA, USA	PRACTICE AREA PPD™ Observational Studies, Thermo Fisher Scientific	CLIENT INTERNAL
---	--	---	---------------------------



Thermo Fisher Scientific
ABS-23

Data Reliability in Retrospective Chart Review Studies: Results and Considerations from a Novel Data Review Methodology

Neil R. Brett,¹ Luke Safarczyk,² Elizabeth Donahue,² Marielle Bassel¹
¹Thermo Fisher Scientific, Montreal, Canada; ²Thermo Fisher Scientific, Waltham, MA, USA

Background

- Retrospective historical cohort studies (chart review studies) are generally efficient designs to generate longitudinal observational follow-up data.¹
- For data review and abstraction verification in clinical studies, source data verification (SDV) may be thought of as the gold-standard.²
- However, the potential use of SDV in chart review studies may face challenges in implementation, site interest, data protection, and timelines.
- Thus, chart review studies generally use remote data review (without source documents); however, this method can result in data reliability/accuracy issues for studies.
- There is a need for a data review methodology that is realistic for efficient implementation in the chart review context, yet still with additional rigor compared with traditional remote data review.

Objective

- The objective was to describe an alternative method for data quality review and highlight key considerations.

Methods

Figure 1. Three historical cohort studies via retrospective chart review

- We performed a review of the three historical cohort studies via retrospective chart review.
- The summary of study status and key evaluations conducted are presented in the Figure 1.

Results

Table 1. Overview of Study Types, Populations and Key Outcomes

Site-based or De-centralized	Study Population and N	Therapeutic Area	Type of Chart Review Study	Types of Data Collected
Site-based	Adults, n=227	Gastroenterology	Natural History Study	Medical history, treatment patterns, clinical outcomes
Site-based	Pediatrics, n=30	Rare Disease	Label Expansion	Patient demographics, medical history, initiation of treatment, clinical outcomes
Site-based	Adults, n=18	Immunology	Drug Utilization	Patient demographics, medical history, treatment patterns including adherence, laboratory tests

Figure 2. Data Review Process Overview of Repeat Abstraction of Key Variables

Abbreviations: CRF = case report form; CRO = contract research organization; EDC = electronic data capture; QC = quality control

Figure 3. Details and Discrepancies for Re-abstraction of Variables

Number of pages/variables of repeat abstraction per study:		
Study 1: 1 page including 13 variables	Study 2: 3 pages including 22 variables	Study 3: 7 pages including 61 variables

Number of patients/sites per study included in re-abstraction:		
Study 1: 7 sites, re-abstraction for 1 patient per site (7 total patients)	Study 2: 17 sites, re-abstraction for all patients (30 total patients)	Study 3: 4 sites, re-abstraction for 1 patient per site (5 total patients)

Re-abstraction discrepancies:		
Study 1: 4 patients with 0 discrepancies, 1 patient with 1%-9% discrepancies and 2 patients with 30%-40% discrepancies	Study 2: 26 patients with 10%-39% discrepancies, 4 patients with >40% discrepancies	Study 3: 0 discrepancies for all patients

Figure 4. Considerations for Repeat Abstraction of Key Variables Data Review

- Study-specific quality control (QC) plans should include:
 - Specific Variables for Re-abstraction:** Focus on variables related to the primary objective; include variables that are complex or prone to misinterpretation; consider variables that have shown high variability or errors in past similar studies.
 - Numbers of Patients:** Determine based on the study's purpose and data complexity; for regulatory studies or those with complex data, include a larger sample size for repeat abstraction purposes; include the first patient(s) abstracted by each site staff member to catch early entry errors.
 - Define Handling of Discrepancies:** Include steps for data correction, site re-training and potential re-entry of data; establish thresholds for acceptance error rates and actions to take if exceeded.
- Additional Considerations:**
 - Timing of Repeat Abstraction:** Initiate repeat abstraction soon after first patient data is abstracted to catch and correct errors early; for studies with larger sample sizes, consider one more than one repeat abstraction to ensure ongoing data quality.
 - Resource and Staffing:** Ensure sites have adequate resources and staff to perform initial and repeat abstraction; avoid having the same person perform initial and repeat abstractions to reduce bias.
 - Study Timelines:** Integrate repeat abstraction activities into the overall study timeline; allow sufficient time for data review, discrepancy resolution, and any necessary re-training.
 - Training and Support:** Provide comprehensive training for site staff on the abstraction process; offer ongoing support and resources to address questions and issues as they arise.
 - Compensation:** Include repeat abstraction fees into the budget to compensate sites for the work performed; if there are study budget constraints, consider reasonably minimizing the number of pages/variables and the number of patients for re-abstraction.
 - Documentation and Reporting:** Maintain detailed records of all abstraction activities, discrepancies, and resolutions; report on data quality metrics and any corrective actions taken.

Study types that may be suitable for this methodology: Retrospective studies where on-site monitoring or source data verification may not occur.

Conclusions

- The data review methodology of repeat abstraction of key variables demonstrates the importance of additional data QC in chart review studies, which enhances the reliability and accuracy of data abstraction, leading to more robust study outcomes.
- This cost-efficient methodology should be customized to each study based on study design, outcomes, timelines, and other relevant factors.
- Studies looking to implement this data review methodology should tailor it to be study-specific by carefully considering data abstraction complexity, study timelines, site burden, and purpose of the study.

References

- Stang KG, et al. Clinical Research Methods in the Rheumatic Diseases. In: Firestein GS, Budd RC, Gabriel SE, et al., eds. Kelley and Firestein's Textbook of Rheumatology. 10th ed. Vol 1. Elsevier; 2017:478-488.
- Anderson JR, et al. Br J Clin Pharmacol. 2015 Apr;79(4):660-8.
- Hansen M, et al. J Clin Pharm Sci. 2024 May 21;1(1):101.

Disclosures

All authors are employees of PPD™ Observational Studies, Thermo Fisher Scientific. Funding provided by Thermo Fisher Scientific.

Acknowledgments

Editorial and graphic design support were provided by Michael Grouss and Richard Leason of Thermo Fisher Scientific.

Presented at the GetReal Conference 2025 | 1-2 October 2025 | Utrecht, The Netherlands

© 2025 Thermo Fisher Scientific Inc. All rights reserved. All trademarks are the property of Thermo Fisher Scientific and its subsidiaries unless otherwise specified. This information is not intended to encourage use of these products in any manner that might infringe the intellectual property rights of others.

GetReal 2025

1–2 October 2025 | Utrecht, The Netherlands

Poster Title

Cost-effectiveness of Radiotherapy in Uterine Serous Carcinoma (USC): A Real-world Study

Objective

- To describe the cost-effectiveness of adjuvant radiotherapy in USC patients treated in the real-world setting.

<p>AUTHORS</p> <p>Eleferios (Pierre) Samartzis¹ Neil Brett³ Lucy Gilbert² Marielle Bassel³ Vincent McCarty² John Sampalis^{2,3}</p>	<p>AFFILIATION</p> <p>¹University Hospital Zürich, Zürich, Switzerland ²Thermo Fisher Scientific, Montreal, Canada ³McGill University, Montreal, Canada</p>	<p>PRACTICE AREA</p> <p>PPD™ Observational Studies, Thermo Fisher Scientific</p>	<p>CLIENT</p> <p>INTERNAL University Hospital, Zürich McGill University, Montreal</p>
---	---	---	--

Cost-effectiveness of Radiotherapy in Uterine Serous Carcinoma (USC): A Real-world Study



ABS-25

Eleferios (Pierre) Samartzis,¹ Lucy Gilbert,² Vincent McCarty,² Neil Brett,³ Marielle Bassel,³ John Sampalis^{2,3}
¹University Hospital Zürich, Zürich, Switzerland; ²Thermo Fisher Scientific, Montreal, Canada; ³McGill University, Montreal, Canada

Background

- Approximately 10% of all endometrial cancers (ECs) are uterine serous (USC).¹
- Distribution of stage at diagnosis is predominantly stage I (40%), followed by stage II (30%), stage III (20%), and stage IV (10%).²
- However, due to its aggressive nature, USC EC is associated with high recurrence and poor prognosis, and it accounts for 40% of EC-related deaths.²
- Treatment of EC may consist of multiple components.
 - Adjuvant chemotherapy (C) with carboplatin + paclitaxel is always recommended.
 - Surgery recommendations depend on diagnosis (e.g., total hysterectomy, bilateral salpingo-oophorectomy, pelvic and para-aortic lymph node dissection, omental biopsy and pelvic washings for staging and detection of spread cancer).
 - The benefit of adjuvant radiotherapy (vaginal brachytherapy or external beam radiotherapy [EBRT]) is currently unclear.
- Since the potential benefit and optimal timing of adjuvant radiotherapy is unclear additional real-world evidence is needed to understand its effectiveness.

Objective

- To describe the cost-effectiveness of adjuvant radiotherapy in USC patients treated in the real-world setting.

Methods

- Retrospective observational cohort study
- Patients with USC receiving:
 - C or C + adjuvant external beam radiotherapy (EBRT)
- All patients were treated at the McGill University Health Center (MUHC) between 2008 and 2023.
- Patient treatment characteristics and outcomes were ascertained from the MUHC Electronic Health Records and the MUHC Gyno-Oncology Database.
- Costs for radiotherapy were assumed to be 2025 average costs in the US.
- Incremental cost-effectiveness ratios (ICERs) were estimated per year of progression-free (PFS) and overall survival (OS) benefit.

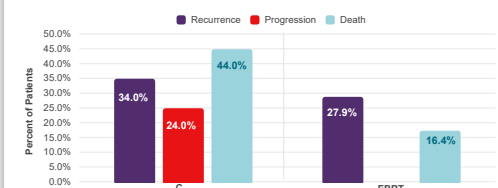
Results

Table 1. Patient characteristics and outcomes

	Treatment Group		P-value
	C (n=50)	C+EBRT (n=61)	
Mean (SD) age, years	67.30 (9.07)	69.40 (8.11)	0.219
Stage I	13 (26.0%)	29 (47.5%)	
Stage II	0 (0.0%)	13 (21.3%)	
Stage III	13 (26.0%)	16 (26.2%)	<0.001
Stage IV	24 (48.0%)	1 (1.6%)	
Preop albumin < 3.5	6 (12.0%)	1 (1.6%)	0.028
Cytology Malignant	23 (46.0%)	9 (14.8%)	<0.001
MMR Status Deficient	1 (2.0%)	7 (11.5%)	0.046

Abbreviations: C = chemotherapy; EBRT = external beam radiotherapy; MMR = mismatch repair status

Figure 1. Proportion of patients with clinical outcomes



	Treatment Group		P-value
	C (n=50)	C+EBRT (n=61)	
Recurrence	17 (34.0%)	17 (27.9%)	0.560
Progression	12 (24.0%)	0 (0.0%)	<0.001
Death	22 (44.0%)	10 (16.4%)	0.005

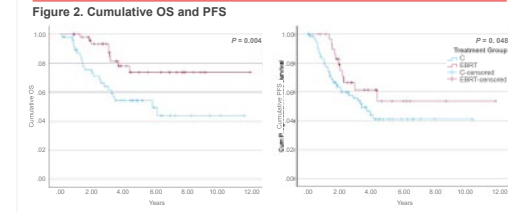
Abbreviations: C = chemotherapy; EBRT = external beam radiotherapy

References

- Egami G, et al. *Gynecol Oncol*. 2021;152(1):226-34.
- Ball A, et al. *Journal of Obstetrics and Gynaecology Canada*. 2014;36(12):1085-92.

Results (cont.)

Cumulative OS and PFS



Abbreviations: C = chemotherapy; EBRT = external beam radiotherapy; PFS = progression-free survival

Regression and Cost-effectiveness

Table 2. Cox regression analyses understanding how covariates associate with clinical outcomes

	OS				PFS			
	P-value	HR	Lower	Upper	P-value	HR	Lower	Upper
EBRT vs C	0.242	0.571	0.223	1.459	0.141	0.584	0.285	1.196
Stage (III-IV) vs (I-II)	0.040	2.726	1.046	7.109	0.004	2.315	1.463	6.053
Pre-Op Albumin <= 3.5	0.685	1.362	0.306	6.056	0.926	0.934	0.218	4.004
MMR Deficient	0.978	0.000	0.000	—	0.389	0.531	0.126	2.242

MMR = Mismatch Repair; Abbreviations: C = chemotherapy; EBRT = external beam radiotherapy; HR = hazard ratio; OS = overall survival; PFS = progression free survival

Table 3. Describing cost-effectiveness of treatments

Treatment	N	Cost per EBRT	Total Cost for EBRT	Total Survival Years	Total PFS Years
EBRT	100	\$27,812.00	\$2,781,200.00	998.7	745.8
C	100	\$0.00	\$0.00	681.9	536.4
Difference (EBRT-C)				316.8	209.4
Incremental cost per year				\$8,779.04	\$13,281.76

Abbreviations: C = chemotherapy; EBRT = external beam radiotherapy; PFS = progression free survival

Discussion

- The results of this study have shown that adjuvant radiation therapy provides clinically important benefits in terms of OS and PFS for patients with USC EC.
- Adjuvant radiation therapy had an acceptable ICER for both OS and PFS.
- Limitations**
 - This was a single-site study conducted in a tertiary center that is highly specialized in the treatment of gynecological cancers.
 - Therefore, the study population may not be representative of the general patient population with USC EC.
 - This study was conducted in Canada, where a universal, publicly funded healthcare system is in effect.
 - Results may be different in non-public/universal healthcare systems, where access to care may be a barrier to receiving radiation therapy.
 - The study was conducted prior to increased use of targeted and immunotherapy.
 - Thus, results may vary among inpatients treated with advanced therapies.

Conclusions

- Adjuvant radiation therapy (EBRT) may be beneficial and cost-effective in a universal healthcare system.
- Further studies are needed to confirm these results in other health care systems.
- Real-world studies are required to conduct cost-effectiveness assessments of treatments.
- Evidence from these studies will drive decisions regarding optimal use of high-cost treatments for rare diseases and cancers.

Disclosures

NR, NB, and JB are employees of PPD™ Observational Studies, Thermo Fisher Scientific. Funding for this project was provided by Thermo Fisher Scientific.

Acknowledgements

Editorial and graphic design support were provided by Caroline Cole and Heather Neuhoff of Thermo Fisher Scientific.



GetReal 2025

1–2 October 2025 | Utrecht, The Netherlands

Poster Title

A Real-world-derived Algorithm to Improve the Cost-effectiveness of External Beam Radiotherapy in Uterine Serous Carcinoma

Objective

- The objective was to develop and evaluate an empirical algorithm aimed at identifying patients with USC EC who would benefit from adjuvant radiotherapy.

<p>AUTHORS</p> <p>Eleferios (Pierre) Samartzis¹ Neil Brett³</p> <p>Lucy Gilbert² Marielle Bassel³</p> <p>Vincent McCarty² John Sampalis^{2,3}</p>	<p>AFFILIATION</p> <p>¹University Hospital Zürich, Zürich, Switzerland</p> <p>²Thermo Fisher Scientific, Montreal, Canada</p> <p>³McGill University, Montreal, Canada</p>	<p>PRACTICE AREA</p> <p>PPD™ Observational Studies, Thermo Fisher Scientific</p>	<p>CLIENT</p> <p>INTERNAL</p> <p>University Hospital, Zürich</p> <p>McGill University, Montreal</p>
---	---	---	--

A Real-world-derived Algorithm to Improve the Cost-effectiveness of External Beam Radiotherapy in Uterine Serous Carcinoma



ABS-26

Eleferios (Pierre) Samartzis,¹ Lucy Gilbert,² Vincent McCarty,² Neil R. Brett,³ Marielle Bassel,³ John Sampalis^{2,3}

¹University Hospital Zürich, Zürich, Switzerland; ²Thermo Fisher Scientific, Montreal, Canada; ³McGill University, Montreal, Canada

Background

- Approximately 10% of all endometrial cancers (ECs) are uterine serous (USC)¹
- Distribution of stage at diagnosis is predominantly stage I (40%), followed by stage II (30%), stage III (20%), and stage IV (10%)²
- However, due to its aggressive nature, USC EC is associated with high recurrence and poor prognosis, and it accounts for 40% of EC-related deaths³
- Treatment of EC may consist of multiple components.
 - Adjuvant chemotherapy (C) with carboplatin + paclitaxel is always recommended.
 - Surgery recommendations depend on diagnosis (e.g., total hysterectomy, bilateral salpingo-oophorectomy, pelvic and para-aortic lymph node dissection, omental biopsy and pelvic washings for staging and detection of spread cancer).
 - The benefit of adjuvant radiotherapy (vaginal brachytherapy or external beam radiotherapy [EBRT]) is currently unclear.
- Since the potential benefit and optimal timing of adjuvant radiotherapy is unclear, understanding its effectiveness and identifying patients who would benefit from EBRT would optimize the utility of this treatment.

Objective

- The objective was to develop and evaluate an empirical algorithm aimed at identifying patients with USC EC who would benefit from adjuvant radiotherapy.

Methods

- Retrospective observational cohort study
- Patients with USC receiving:
 - C
 - C + adjuvant EBRT
- All patients were treated at the McGill University Health Center (MUHC) between 2008 and 2023.
- Patient treatment characteristics and outcomes were ascertained from the MUHC Electronic Health Records and the MUHC Gyno-Oncology Database.
- Multi-variate logistic regression was used to identify predictors of progression or recurrence among those treated with EBRT.
- A receiver operating curve analysis was used to identify three groups with high, moderate, and low potential benefit from EBRT.

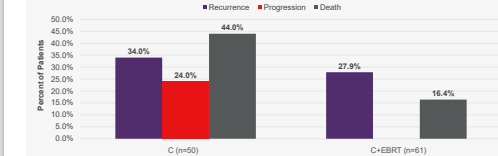
Results

Table 1. Patient characteristics

	Treatment Group		P-value
	C (n=50)	C+EBRT (n=61)	
Mean (SD) age, years	67.30 (9.07)	69.40 (8.11)	0.219
Stage I	13 (26.0%)	29 (47.5%)	
Stage II	0 (0.0%)	13 (21.3%)	<0.001
Stage III	13 (26.0%)	18 (29.5%)	
Stage IV	24 (48.0%)	1 (1.6%)	
PR+	27 (54.0%)	40 (65.6%)	0.028
Cytology			
Malignant	23 (46.0%)	9 (14.8%)	<0.001

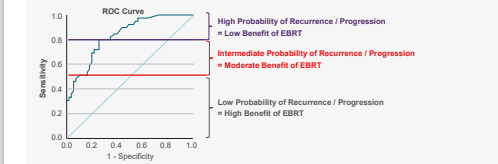
Abbreviations: C = chemotherapy; EBRT = external beam radiotherapy; PR = progesterone receptor

Figure 1. Proportion of patients with outcomes, stratified by treatment



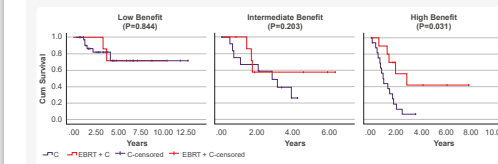
Abbreviations: C = chemotherapy; EBRT = external beam radiotherapy; PRISMA = Preferred Reporting Items for Systematic Reviews and Meta-Analyses

Figure 2. Classification of EBRT benefit categories (P<0.001)



Abbreviations: EBRT = external beam radiotherapy; ROC = receiver operating characteristic

Figure 3. Progression-free survival by benefit categories

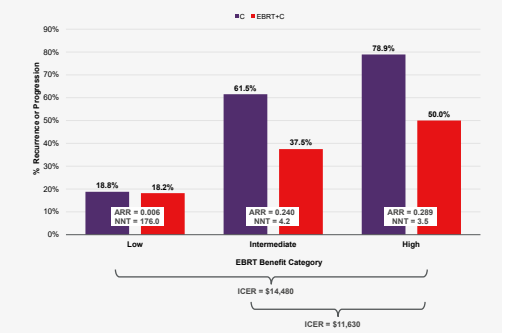


Abbreviations: C = chemotherapy; EBRT = external beam radiotherapy

References

- Bogren G, et al. Uterine serous carcinoma. *Gynecol Oncol*. 2021 Jul;162(1):226-234.
- Bal A, et al. The importance of surgical staging in women with uterine serous carcinoma: experience in a single institution reveals a survival benefit. *J Ovarian Oncol*. 2014 Dec;36(12):1985-1992.

Figure 4. Recurrence or progression by EBRT benefit category



Abbreviations: ARR = absolute risk reduction; EBRT = external beam radiotherapy; ICER = incremental cost effectiveness ratio / progression-free survival (PFS) year gained; NNT = numbers needed to treat

Discussion

- The real-world-derived algorithm for the selection of patients with USC ECs to be treated with EBRT has the potential to:
 - Improve the number needed to treat (NNT)
 - Reduce overall costs and improve cost-effectiveness
 - Prevent potentially non-beneficial EBRT in 33/61 (54.1%) of the patients
 - With associated impact on quality of life and treatment-related adverse events

Limitations

- This was a single-site study conducted in a tertiary center that is highly specialized in the treatment of gynecological cancers. Therefore, the study population may not be representative of the general patient population with USC EC.
- This study was conducted in Canada, where a universal, publicly funded healthcare system is in effect. Results may be different in non-public/universal healthcare systems, where access to care may be a barrier to receiving radiation therapy.
- The study was conducted prior to increased use of targeted and immunotherapy.
- Thus, results may vary among inpatients treated with advanced therapies.

Conclusions

- Adjuvant radiation therapy (EBRT) may be beneficial and cost-effective in a universal healthcare system.
- Using a real-world-derived algorithm to select patients who are most likely to benefit from EBRT may reduce overall costs.

Disclosures

NRB, MB, and JS are employees of PPD™ Observational Studies, Thermo Fisher Scientific. Funding for this poster was provided by Thermo Fisher Scientific. Editorial and graphic design support were provided by Caroline Cole and Richard Leason of Thermo Fisher Scientific.



GetReal 2025

1–2 October 2025 | Utrecht, The Netherlands


Poster Title

Comparing Randomized Clinical Trials to Real-world Studies Evaluating the Effectiveness of a bDMARD in the Management of Crohn's Disease

Objective

- To compare patient profiles and outcomes between RCTs and RWS in the evaluation of a bDMARD treatment for CD.

<p>AUTHORS Vincent McCarty¹ Paige Kostoulas¹ Neil R. Brett² Marielle Bassel² John Sampalis^{1,2}</p>	<p>AFFILIATION ¹McGill University, Montreal, Canada ²Thermo Fisher Scientific, Montreal, Canada</p>	<p>PRACTICE AREA PPD™ Observational Studies, Thermo Fisher Scientific</p>	<p>CLIENT INTERNAL McGill University, Montreal</p>
--	--	--	---


ABS-27

Comparing Randomized Clinical Trials to Real-world Studies Evaluating the Effectiveness of a bDMARD in the Management of Crohn's Disease

Vincent McCarty¹, Paige Kostoulas¹, Neil R. Brett², Marielle Bassel², John Sampalis^{1,2}
¹McGill University, Montreal, Canada; ² Thermo Fisher Scientific, Montreal, Canada

Background


- Recent studies in inflammatory bowel disease (IBD) highlight differences between randomized controlled trial (RCT) populations and populations in RWS, including age, comorbidities, disease severity, and other parameters that could affect treatment responses.^{1,2}
- Because of this, real-world studies (RWS) are becoming increasingly impactful in supporting evidence of treatment effectiveness and safety due to better generalizability to the target population.
- There are currently multiple biologic disease-modifying antirheumatic drugs (bDMARDs) approved for Crohn's disease (CD). However, it is not known if RCTs of bDMARDs have similar patient populations and outcomes to RWS of these treatments.

Objective

- To compare patient profiles and outcomes between RCTs and RWS in the evaluation of a bDMARD treatment for CD.

Methods

Figure 1. Study design
 Data on patients treated for CD with the same bDMARD (identified via PubMed database search) were extracted from:



Abbreviations: CD = Crohn's disease; bDMARD = biologic disease-modifying antirheumatic Drug; RCT = randomized controlled trial; RWS = real-world studies

Statistics

- Categorical variables (sex, prior anti-tumor necrosis factor [TNF] exposure, immunosuppressant and corticosteroid use, Montreal Classification, clinical remission rates) were compared between RWS and RCT cohorts using two-sample z-tests for proportions. Statistical significance for categorical comparisons was defined as p<0.05.
- Continuous variables (age, disease duration, weight, C-reactive protein [CRP], fecal calprotectin) were summarized descriptively.

Results

- In total, 1,176 patients in RCTs and 5,267 in RWS were treated with the bDMARD.
- On average, RWS patients were older, had longer disease duration, and had higher CRP, among other differences, compared with RCT patients (Table 1).

Table 1. Baseline characteristics

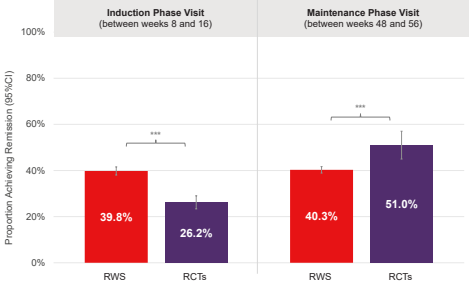
Parameter	19 Large-scale RWS		3 Pivotal Phase 3 RCTs		P-value	Percentage Difference, RWS vs RCT
	Proportion of Patients	n	Proportion of Patients	n		
Male sex (%)	45.8%	4,933	43.1%	1,176	0.0913	6.3%
Prior TNF (%)	84.2%	4,593	65.9%	1,176	<0.001	27.8%
Immunosuppressants (%)	32.6%	4,145	34.4%	1,176	0.2675	-5.0%
Corticosteroids (%)	39.3%	3,876	44.5%	1,176	0.0015	-11.7%
Montreal Classification						
L1 (%)	27.5%	4,339	18.9%	1,176	<0.001	45.8%
L2 (%)	17.6%	4,462	18.3%	1,176	0.5642	-3.9%
L3 (%)	52.1%	4,462	62.5%	1,176	<0.001	-16.6%
L4 (%)	10.8%	2,663	17.9%	1,176	<0.001	-39.8%
P (%)	31.6%	4,462	35.7%	1,176	0.0076	-11.5%
	Mean	n	Mean	n		Percent difference, RWS vs RCT
Age (years)	40.4	5,144	38.0	1,176		6.2%
Disease duration (years)	11.7	4,878	10.5	1,176		11.7%
Weight (kg)	69.4	672	70.7	1,176		-1.9%
CRP (mg/L)	9.6	1,582	9.0	1,176		6.9%
Fecal calprotectin (mg/kg)	495.6	1,263	504.8	1,176		-1.8%

Abbreviations: CRP = C-reactive protein; RCT = randomized controlled trial; RWS = real-world studies; TNF = tumor necrosis factor. P-values calculated using two-sample z-tests for proportions. Bolded P-values indicate statistical significance, defined as P<0.05. The Montreal Classification of CD defines disease location as L1 (ileum only), L2 (colon only), L3 (ileocolonic), and L4 (upper gastrointestinal/proximal involvement); perianal disease is denoted as "P". Immunosuppressants included azathioprine, mercaptopurine, and methotrexate.

Results (cont.)

- Clinical remission rates during induction were significantly higher in the RWS (39.8%) compared to the RCTs (26.2%; P<0.001). However, clinical remission rates during the maintenance phase were significantly lower in the RWS (40.3%) versus RCTs (51.0%, P<0.001) (Figure 1).

Figure 1. Remission outcomes



Abbreviations: RCT = randomized controlled trial; RWS = real-world studies
 *** = P<0.001. P-values calculated using two-sample z-tests for proportions. Error bars represent 95% CI. Number of patients included in the analysis: from left to right: n=2,895; n=912; n=4,054; n=264.

Discussion

- This comparison of RCTs and RWS evaluating a bDMARD in CD revealed important differences in:
 - Baseline patient characteristics
 - Rates of achieving remission
- The differences observed likely reflect the broader inclusiveness of RWS.
- RCT data suggested continued improvements in remission achievement after induction (>16 weeks).
- The results from the RWS suggested that approximately the same proportion of patients would achieve remission through induction as through the first year, underscoring the importance of early outcomes.
- Limitations of this analysis include:
 - Focus on one bDMARD, meaning generalizability to other bDMARDs is unknown
 - Reliance on aggregated study-level results (rather than patient-level data)
 - Variation among RWS, including the timing of follow-up visits and reporting of baseline characteristics

Conclusions

- The significant differences in patient characteristics and clinical outcomes between RWS and RCTs highlight the necessity of real-world evidence in a comprehensive assessment of marketed and new treatments.
- RWS provide important insights into patient populations and clinical effectiveness that differ from those assessed by RCTs, offering a complementary role in understanding the true therapeutic potential and limitations of treatments in real-world practice.

References


- Ingrascio Y, et al. *Pharmacol Res.* 2024 Feb;205:107074.
- Vieujen S, et al. *Lancet Healthy Longev.* 2022 May;3(5):e356-e366.

Disclosures

NRB, MB, and JS are employees of PPD™ Observational Studies, Thermo Fisher Scientific. Funding for this poster was provided by Thermo Fisher Scientific.

Acknowledgments

Editorial and graphic design support was provided by Caroline Cole and Shari Berger of Thermo Fisher Scientific.



Presented at the GetReal Conference 2025 | 1–2 October 2025 | Utrecht, The Netherlands

© 2025 Thermo Fisher Scientific Inc. All rights reserved. All trademarks are the property of Thermo Fisher Scientific and its subsidiaries unless otherwise specified. This information is not intended to encourage use of these products in any manner that might infringe the intellectual property rights of others.

GetReal 2025

1–2 October 2025 | Utrecht, The Netherlands

Poster Title

Comparison of Randomized Controlled Trials (RCT) with Open-label, Single-cohort, Real-world Studies (RWS) in the evaluation of pharmacologic treatment of Non-small Cell Lung Cancer (NSCLC)

Objective

- The aim of this study was to assess the validity of using ECAs derived from real-world studies in the assessment of efficacy of treatments for non-small cell lung cancer (NSCLC).
 - Efficacy outcomes assessed were objective response rate (ORR) and progression-free survival (PFS)

<p>AUTHORS</p> <p>Richard LeBrun Trachy¹</p> <p>Vincent McCarty¹</p> <p>Neil R. Brett²</p> <p>Marielle Bassel²</p> <p>John Sampalis^{1,2}</p>	<p>AFFILIATION</p> <p>¹McGill University, Montreal, Canada</p> <p>²Thermo Fisher Scientific, Montreal, Canada</p>	<p>PRACTICE AREA</p> <p>PPD™ Observational Studies, Thermo Fisher Scientific</p>	<p>CLIENT</p> <p>INTERNAL</p> <p>McGill University, Montreal</p>
--	--	---	---

Comparison of Randomized Controlled Trials (RCT) with Open-label, Single-cohort, Real-world Studies (RWS) in the evaluation of pharmacologic treatment of Non-small Cell Lung Cancer (NSCLC)



ABS-28

Richard LeBrun Trachy,¹ Vincent McCarty,¹ Neil R. Brett,² Marielle Bassel,² John Sampalis^{1,2}
¹McGill University, Montreal, Canada; ²Thermo Fisher Scientific, Montreal, Canada

Background

- Randomized controlled trials (RCT) have been considered the gold-standard in evaluating treatment efficacy.
 - However, in some cases of rare diseases and end-stage cancers, RCTs are not feasible or ethical.
- In recent years, real-world evidence has become increasingly important in the assessment of efficacy and safety of novel and existing treatments, addressing requirements for all stakeholders including regulators, patients, and healthcare providers.
- In many cases, the use of external control arms (ECAs) is an efficient way of conducting comparative analyses and complementing RCTs.
 - Regulatory agencies are increasingly more accepting of ECAs given certain requirements of quality, minimal bias, and comparability with the target or study population are demonstrated.
- There are concerns regarding the interpretation and validity of results from real-world, single-cohort studies as ECAs with respect to validity and representation of results that would be obtained in RCTs.

Objective

- The aim of this study was to assess the validity of using ECAs derived from real-world studies in the assessment of efficacy of treatments for non-small cell lung cancer (NSCLC).
 - Efficacy outcomes assessed were objective response rate (ORR) and progression-free survival (PFS)

Methods

- Literature search (PubMed)
 - ((((Randomized controlled trial) OR RCT)) OR randomized)) AND ((NSCLC) OR non-small cell lung cancer).
 - Records were then excluded based on the criteria, yielding a final set of 385 RCTs.
 - Single-cohort studies assessing any of the 21 treatment protocols used as experimental or control arms in the RCTs were identified.
 - ((((Single-Arm) OR Single Arm)) OR nonrandomized)) AND ((NSCLC) OR non-small cell lung cancer). This search yielded 89 single-cohort studies.
- Finally, the comparable treatment were cross-verified to ensure it was evaluating the same treatment for single-cohort studies and RCTs.
- This resulted in 87 RCTs and 48 single-cohort studies used in the final evaluation for 21 treatments for NSCLC.

Results

Table 1. Summary of Patient Characteristics and Outcomes Across Treatments

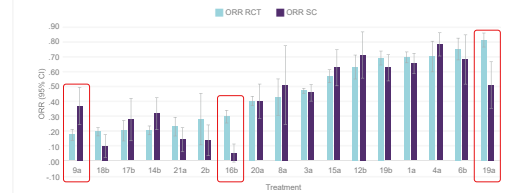
Treatment Number	RCT				Single-cohort Study					
	N	Age (Years) ^a	Male %	ORR ^a	PFS ^a	N	Age (Years) ^a	Male %	PFS ^a	
1a	632	61	38	68	76	576	66	35	50	77
1b	32	67	84	9	8	18	39	33	11	11
2b	22	59	82	27		38	59	68	13	11
3a	3211	62	60	46	65	306	63	75	45	62
4a	75	67	40	69	90	109	66	39	77	84
4a	75	67	40	69	90	109	66	39	77	84
5b	39	61	50	64	39	76	63	59	20	33
6b	137	58	50	74	81	28	54	49	67	76
8a	60	58	93	42	54	10	77		50	20
9a	397	67	61	17	43	53	65	77	36	39
10b	115	54	41		42	227	53	47	55	65
12b	138	60	41	62	65	30	56	44	70	83
14b	699	64	65	29	35	68	61	57	31	38
15a	507	64	37	56	66	355	69	27	62	67
15b	1916	63	58	15	29	326	66	64		11
16b	434	66	57	29	44	43	57	30	5	26
17b	126	61	57	20	38	136	69	64	27	32
18b	765	61	71	19	33	53	61	62	9	21
19a	279	64	36	80	88	36	60	61	50	64
19b	384	60	40	68	78	228	65	30	69	70
20a	945	64	63	39	57	132	68	67	51	60
21a	168	71	60	22	36	66	76	65	14	35
21b	423	64	60	15	38	30	67	97		35

^aWeighted Average
 Abbreviations: ORR = objective response rate; PFS = progression-free survival; RCT = randomized controlled trial

Results (cont.)

- Objective Response Rate (Figure 1)**
- The ORR reported in RCTs and real-world studies was not statistically significant for 14/17 (82.4%) of treatments
 - The red boxes indicate ORR was significantly higher for the RCT in 2/17 (11.7%) and higher for real-world study in 1/17 (5.8%)

Figure 1. ORR by Treatment and Study Type

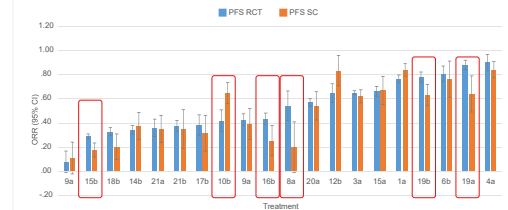


Abbreviations: ORR = objective response rate; RCT = randomized controlled trial; SC = single cohort, first-line treatment (a), subsequent-line treatment (b)

Progression-free Survival (Figure 2)

- PFS was not significantly different for 14/20 (70%) treatments.
- The red boxes indicate PFS was significantly higher in the RCT for 6/20 (30%) treatments.

Figure 2. PFS by Treatment and Study Type



Abbreviations: PFS = progression-free survival; RCT = randomized controlled trial; SC = single cohort, first-line treatment (a), subsequent-line treatment (b)

Discussion

- The results of this study demonstrated that the ORR and PFS observed in RCTs and single-cohort studies utilizing ECAs evaluating the treatment of NSCLC are comparable.

Limitations

- Focus on NSCLC: Generalization to other therapeutic areas requires further assessment.
- Publication Bias: Studies reported in the literature may be biased towards positive results.
- Design Limitations of single-cohort studies:
 - Random selection of patients could not be assessed
 - Timing of single-cohort study relative to the RCT was not assessed
 - Potential for bias by indication
 - Agreement with ECAs may be higher if single-cohort studies were specifically designed as ECAs

Conclusions

- Well designed and carefully conducted single-cohort studies could provide ECAs replacing comparators in RCTs
- When fully randomized RCTs are not feasible or ethical, a real-world ECA would be a viable alternative.
- Real-world studies can enhance and supplement the evaluation of treatments for rare diseases or advanced cancers.

Disclosures

NR, NB, and JB are employees of PPD™ Observational Studies, Thermo Fisher Scientific. Funding for this poster was provided by Thermo Fisher Scientific.

Acknowledgements

Editorial and graphic design support were provided by Michael Gross and Heather Neuhoff of Thermo Fisher Scientific.



GetReal 2025

1–2 October 2025 | Utrecht, The Netherlands

Poster Title

Utility of Real-world External Control Arm (ECA) in the Evaluation of Pharmacologic Treatment of Non-small Cell Lung Cancer (NSCLC)

Objective


- The objective of this study was to compare the results observed in RCTs and real-world studies (RWS) assessing the efficacy of treatments for non-small cell lung cancer (NSCLC).
- The efficacy outcomes assessed were objective response rate (ORR), progression-free survival (PFS) and overall survival (OS).

<p>AUTHORS</p> <p>Richard LeBrun Trachy¹</p> <p>Vincent McCarty¹</p> <p>Neil R. Brett²</p> <p>Marielle Bassel²</p> <p>John Sampalis^{1,2}</p>	<p>AFFILIATION</p> <p>¹McGill University, Montreal, Canada</p> <p>²Thermo Fisher Scientific, Montreal, Canada</p>	<p>PRACTICE AREA</p> <p>PPD™ Observational Studies, Thermo Fisher Scientific</p>	<p>CLIENT</p> <p>INTERNAL</p> <p>McGill University, Montreal</p>
--	--	---	---

Utility of Real-world External Control Arm (ECA) in the Evaluation of Pharmacologic Treatment of Non-small Cell Lung Cancer (NSCLC)

Richard LeBrun Trachy,¹ Vincent McCarty,¹ Neil R. Brett,² Marielle Bassel,² John Sampalis^{1,2}

¹McGill University, Montreal, Canada; ²Thermo Fisher Scientific, Montreal, Canada



ABS-29

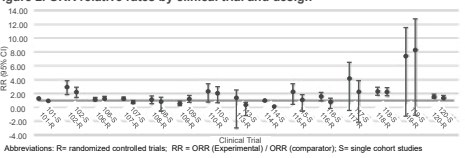
Background

- Randomized controlled trials (RCT) have been considered the gold standard in evaluating treatment efficacy.
- However, in some cases RCTs are not feasible or ethical, like with rare diseases and end-stage cancers.
- In recent years, real-world evidence has become increasingly important in the assessment of the efficacy and safety of novel and existing treatments, addressing requirements for all stakeholders, including regulators, patients, and healthcare providers.
- In many cases, the use of external control arms (ECAs) is an efficient way of conducting comparative analyses and complementing RCTs.
- Regulatory agencies are increasingly more accepting of ECAs given certain requirements of quality, minimal bias, and comparability with the target or study population are demonstrated.
- There are concerns regarding the interpretation and validity of results from real-world, single-cohort (SC) studies with respect to validity and representation of results that would be obtained in RCTs.

Results (cont.)

- For ORR, there was agreement between the RCTs and ECA-RCTs for 9/15 (60%) in **Figure 2**.
- The purple box indicates non-significant different direction (NSDD) in 5/15 (33.3%).
- The red box indicates significant in different direction (SDD) in 1/15 (6.7%).

Figure 2. ORR relative rates by clinical trial and design



Abbreviations: R= randomized controlled trials; RR = ORR (Experimental) / ORR (comparator); S= single cohort studies

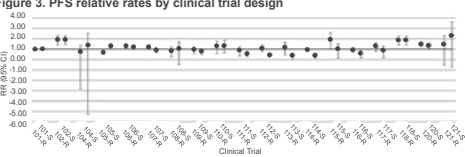
Objective

- The objective of this study was to compare the results observed in RCTs and real-world studies (RWS) assessing the efficacy of treatments for non-small cell lung cancer (NSCLC).
- The efficacy outcomes assessed were objective response rate (ORR), progression-free survival (PFS) and overall survival (OS).

Results (cont.)

- For PFS, there was agreement between the RCTs and ECA-RCTs for 12/15 (80.0%) in **Figure 3**.
- The purple box indicates NSDD in 2/15 (13.3%).
- The red box indicates SDD in 1/15 (6.7%).

Figure 3. PFS relative rates by clinical trial design



Abbreviations: R= randomized controlled trials; RR = ORR (Experimental) / ORR (comparator); S= single cohort studies

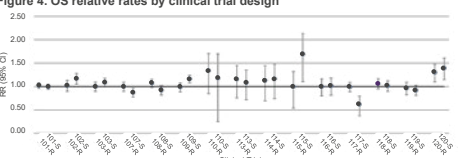
Methods

- The following search string was used for the literature search in PubMed: -((((Randomized controlled trial) OR RCT) OR randomized) AND ((NSCLC) OR non-small cell lung cancer).
- Records were then excluded based on the criteria, yielding a final set of 385 RCTs.
- Single cohort (SC) studies assessing any of the 21 treatment protocols used as experimental or control arms in the RCTs were identified using the following search string: -((((Single-Arm) OR Single Arm)) OR nonrandomized) AND ((NSCLC) OR non-small cell lung cancer). This search yielded 89 SC studies.
- Finally, the comparable treatment was cross-verified to ensure the same treatment was evaluated for both SC and RCTs.
- This resulted in 87 RCTs and 48 SC studies used in the final evaluation of 21 treatments for NSCLC.

Results (cont.)

- For OS, there was agreement (A) between the RCTs and ECA-RCTs for 14/15 (93.3%) in **Figure 4**.
- The red box indicates SDD in 1/15 (6.7%).

Figure 4. OS relative rates by clinical trial design



Abbreviations: R= randomized controlled trials; RR = ORR (Experimental) / ORR (comparator); S= single cohort studies

Discussion

- The results of this study demonstrated that the ORR, progression-free survival (PFS), and OS were comparable in RCTs and SC studies using ECAs and evaluating the treatment of NSCLC.

Conclusions

- Well-designed and carefully conducted SC studies can supplement the evidence provided by RCTs.
- When RCTs are not feasible or ethical, the SC design can be a viable alternative.
- RWS can enhance and supplement the evaluation of treatments for rare diseases or advanced cancers.

Disclosures

Neil, RB, and JB are employees of PPD™ Observational Studies, Thermo Fisher Scientific. Funding for this poster was provided by Thermo Fisher Scientific.

Acknowledgements

Editorial and graphic design support was provided by Caroline Cole and Marissa Carter of Thermo Fisher Scientific.

Presented at the GetReal Conference 2025 | 1–2 October 2025 | Utrecht, The Netherlands

ppd



ObesityWeek

4–7 November 2025 | Atlanta, GA, USA

Posters

Guidance for Obesity Management: What About the Patient's Experience?



Poster Title

Guidance for Obesity Management: What About the Patient’s Experience?

Objective

- To examine the inclusion of qualitative evidence in approval packages and product labels for obesity management medications (OMMs).

<p>AUTHORS Hayley Karn¹ Ana C S Liberato¹ Carla Dias-Barbosa²</p>	<p>AFFILIATION ¹Thermo Fisher Scientific, London, UK ²Thermo Fisher Scientific, Lyon, France</p>	<p>PRACTICE AREA PPD™ Evidera™ Patient-Centered Research, Thermo Fisher Scientific</p>	<p>CLIENT INTERNAL</p>
---	---	---	-----------------------------------



Guidance for Obesity Management: What About the Patient’s Experience?

Hayley Karn¹, Ana C S Liberato¹, Carla Dias-Barbosa²
¹Thermo Fisher Scientific, London, UK; ²Thermo Fisher Scientific, Lyon, France

Background

- The emergence of novel pharmacotherapies has revolutionized the treatment of obesity.
- The 2025 US FDA draft guidance for industry, titled “Obesity and Overweight: Developing Drugs and Biological Products for Weight Reduction”¹ focuses on the design of trials to demonstrate the effectiveness of drugs on sustained weight reduction. However, this lacks consideration of the physical, emotional, and lifestyle impacts experienced by patients with obesity.
- The FDA recognizes the value of qualitative interviews in establishing what is important to patients during drug development.²
- Qualitative in-trial interviews with clinical trial participants have emerged as a leading evidence-generation tool to support the approval of innovative therapies. These interviews offer a unique opportunity to elevate the patient voice by capturing in-depth patient experience data on meaningful treatment experiences and can support the interpretation of clinical outcomes assessment (COA) and other clinical trial efficacy endpoints.

Objective

- To examine the inclusion of qualitative evidence in approval packages and product labels for obesity management medications (OMMs)

Methods

- To assess the inclusion of qualitative data in regulatory submissions we:
 - Identified FDA-approved OMMs for obesity using the Center for Drug Evaluation and Research (CDER) database³
 - Reviewed the corresponding drug approval packages and product labels

Results

- Six OMMs have been FDA approved for obesity and/or overweight between 1999 and 2025 (see **Table 1** for approval and last label review dates): orlistat, phentermine and topiramate, bupropion and naltrexone, liraglutide, semaglutide, and tirzepatide.
- Three OMMs included in-trial and/or other qualitative data as part of the FDA clinical/medical or other review documents.
- No OMM applications included in-trial interview data in the product label.
- Feedback from the FDA on these OMM submissions included recommendations to sponsors for submitting qualitative evidence as part of drug approval packages and the importance of well-designed in-trial research:
 - Engage FDA early on plans for qualitative research
 - Provide comprehensive documentation for early and effective review of qualitative evidence
 - Specify and define patient-relevant functional impacts
 - Use patient interviews to support anchor-based analyses, content validity, and meaningful change
 - Use exit interviews to ask patients about satisfaction and drug effectiveness
 - Involve patients in qualitative research that are representative of the clinical trial population

Results (cont.)

Table 1. FDA Review of OMM Submissions Containing In-Trial and/ or Qualitative Interview Data (N=6)

Active Ingredient/ Generic Drug Name (Brand Name), Company	Drug Class	FDA Approval Date and Latest Label Review Date	Qualitative Data Included	Type of FDA Review Document	Type of Qualitative Research and Objective	FDA Comments Regarding Qualitative Research
Orlistat (Xenical/Alli) Cheplapharm/GSK plc	Gastric and pancreatic lipase inhibitor	FDA approval: 23 Apr 1999 Latest label review: 17 Nov 2022	• Xenical: No • All: Yes	Medical review	Type: In-trial longitudinal interviews Objective: Assess subject-perceived study treatment effectiveness and satisfaction	• Satisfaction and drug effectiveness questions should have been reserved for the final interview.
Phentermine and topiramate (Qsymia) Vivus	Norepinephrine agonist/GABA agonist, glutamate antagonist	FDA approval: 7 Jul 2012 Latest label review: 19 Nov 2024	No	-	-	-
Bupropion hydrochloride and naltrexone hydrochloride (Contrave) Orexigen Therapeutics, Inc.	Naltrexone/bupropion: Opioid receptor antagonist/ dopamine and norepinephrine reuptake inhibitor	FDA approval: 10 Sep 2014 Latest label review: 1 Apr 2022	Yes	Medical review	Type: CE Objective: Evaluate content validity of the IWQOL-Lite	• Qualitative evidence failed to demonstrate the IWQOL-Lite was fit-for-purpose. • Interviewed patients may not be representative of clinical trial population. • Inadequate documentation of content validity. • No submission of qualitative protocols, interview guides, transcripts or publications.
Liraglutide (Saxenda) Novo Nordisk	GLP-1 receptor agonist	FDA approval: 23 Dec 2014 Latest label review: 1 Nov 2024	No	-	-	-
Semaglutide (Wegovy) Novo Nordisk	GLP-1 receptor agonist	FDA approval: 4 Jun 2021 Latest label review: 27 Nov 2024	No	-	-	-
Tirzepatide (Zepbound) Eli Lilly	Dual GLP-1 receptor and GIP receptor agonist	FDA approval: 8 Nov 2023 Latest label review: 2 Jul 2025	Yes	Other reviews and clinical review	Type: CE and CI Objective: 1. Explore impacts of obesity on PF 2. Evaluate content validity of the SF- 36v2 3. Determine MIC on SF-36v2 PF	• Interview results did not inform meaningful change for the SF- 36v2 PF domain score. • Importance of Agency review and comment on qualitative research plans early in the development program • Recommended interviews to support anchor-based analyses, content validity and MIC. • Specify and define functional impacts relevant and important to patients.

Abbreviations: CE= concept elicitation; CI= cognitive interviews; GABA = gamma-aminobutyric acid; GIP = glucose-dependent insulinotropic polypeptide; GLP-1 = glucagon-like peptide-1; IWQOL-Lite = Impact of Weight on Quality of Life-Lite; MIC= minimally importance change; PF = physical functioning; SF-36v2 = Short Form Health Survey version-2.0

Disclosures

HK, ACSL, and CDB are employees of PPD™ Evidera™ Patient-Centered Research, Thermo Fisher Scientific. Funding for this poster was provided by Thermo Fisher Scientific.

Acknowledgments

Editorial and graphic design support were provided by Caroline Cole and Kawthar Nakayima of Thermo Fisher Scientific.

Strengths

- To our knowledge, this is the first review of qualitative evidence in OMM regulatory submissions.
- This review has identified a gap in the FDA drug evaluation and approval process for OMMs: the focus for approvals remains on quantitative data, i.e. weight loss percentages, as evidenced in the recent FDA draft guidance for industry, with less focus on the patient voice.

Limitations

- The review was limited to approved OMMs only.
- The review focused only on the US regulatory landscape (FDA); future research should also evaluate the European landscape (EMA submissions).

Conclusions

- The FDA recognizes the value of qualitative data to:
 - Capture patients’ experiences with OMMs.
 - Support the content validity and interpretability of COA endpoints.
- Whilst the 2022 Patient-Focused Drug Development (PFDD) guidance encourages the inclusion of qualitative evidence;⁵
 - The FDA does not formally require it in submission packages.
 - There appears to have been a lack of opportunity to date for sponsors to collect rich qualitative data to complement COA-based endpoints.
- There is a need for sponsors to design more patient-centered clinical trials integrating the patient voice.
- The FDA recommends for sponsors of OMMs to:
 - Engage the FDA early
 - Provide comprehensive qualitative evidence
 - Specify and define functional impacts
 - Use patient interviews to support anchor-based analyses, content validity and meaningful change
 - Use exit interviews to ask patients about satisfaction and drug effectiveness
 - Ensure qualitative samples represent the clinical trial population
- Well-designed in-trial qualitative studies can help address FDA concerns regarding the interpretation of COA endpoints.
- Collecting rich qualitative data may also reduce the need for sponsors to collect additional real-world evidence research often required to supplement clinical trial data.

References

- US FDA. Obesity and Overweight: Developing Drugs and Biological Products for Weight Reduction. Accessed 10 March 2025. <https://www.fda.gov/media/71252/download>.
- US FDA. Patient-Focused Drug Development: Methods to Identify What Is Important to Patients Guidance for Industry, Food and Drug Administration Staff, and Other Stakeholders. Accessed 10 March 2025. <https://www.fda.gov/media/131230/download>.
- US FDA. Drug Approvals and Databases. Accessed 10 March 2025. <https://www.fda.gov/drugs/development-approval-process-drugs/drug-approvals-and-databases>

The Professional Society for Health Economics and Outcomes Research (ISPOR EU)

9–12 November 2025 | Glasgow, Scotland, UK

Posters

Exploring the Promise of Generative Artificial Intelligence (AI) for Coding and Analysing Qualitative Patient Data: Results from a Pilot Study in Non-Hodgkin's Lymphoma

Evidence of Mapping Malpractice? A Review of Mapping Algorithm Usage in NICE HealthTechnology Appraisals

From Theory to Practice: Guidance for Identifying and Selecting Mapping Algorithms for Health Economic Models

Anchors Away: Navigating Unanchored Indirect Comparisons with Multilevel Unanchored Meta-regression

An Exploration of Implementing Herd Immunity Into Static Models for Economic Evaluation of Vaccines

Using a Large Language Model (LLM) for Data Extraction of Studies: Learnings from a Targeted Literature Review (TLR) in Non-small Cell Lung Cancer (NSCLC)

Real-world Evidence on the Broader Benefits of GLP-1 RA-based Therapies: A Global Targeted Literature Review

Millennial Medical Record Data Profile: A Japanese Electronic Medical Records Database Utilising Unstructured Data for Lung Cancer Research

AI-Assisted Time-to-Event Projection: A Case Study and Broader Potential

Have Delays in Vaccine Market Access Improved in Europe and the US Over the Last Decade?

NICE Has Embraced Single-arm Trials in Technology Appraisals, but Where is the Guidance?

Comparing Alternative Extrapolation Methods Using Standard Partitioned Survival Model Functionality in the Presence of Converging Survival Data: A Case Study in Renal Cell Carcinoma

Are Clinicians Ready For A More Environmental-friendly Health Care? A Clinician Survey

Navigating Health Technology Assessment Requirements: The Current Landscape of Alzheimer's Disease Modifying Treatments

A Layered Approach to Reducing Hallucinations in LLMs for Clinical Data

Identifying Risk Profiles for Early Treatment Discontinuation in Geographic Atrophy Using Machine Learning and SHAP Clustering

Uncovering Patient Narratives of Opioid Use and Recovery Using Large Language Models for Topic and Emotion Analysis of Social Media

Enhancing Migraine Preference Research: Recommendations from a Systematic Review of Preference Studies

Integrating Real-World Evidence into Oncology Health Technology Assessment Submissions: Recent EU Examples

Evaluating High-cost Gene Therapies for Non-cancer Conditions: Insights From NICE's Standard Appraisal Process

Enhancing Targeted Literature Reviews (TLRs) with Artificial Intelligence (AI): A Methodological Approach for Conducting Efficient Targeted Searches

Navigating Evidence Requirements for Joint Clinical Assessment (JCA): Practical Considerations for Systematic Literature Reviews (SLRs) and Indirect Treatment Comparisons (ITCs)

Cost Implications of Declining MMR Coverage in England: Modelling NHS Burden and the Value of Catch-Up Vaccination



Poster Title

Exploring the Promise of Generative Artificial Intelligence (AI) for Coding and Analysing Qualitative Patient Data: Results from a Pilot Study in Non-Hodgkin's Lymphoma

Objective

- Evaluate how AI performs compared with human tasks and processes.
- Provide recommendations for AI use-cases that can meet regulatory standards, using both primary and secondary analysis of qualitative data.

AUTHORS Karen Bailey ¹ Anne Skalicky ² Carla Dias Barbosa ³ Sonya Stanczyk ² Paul Cordero ⁴	AFFILIATION ¹ Thermo Fisher Scientific, London, UK ² Thermo Fisher Scientific, Waltham, MA, USA ³ Thermo Fisher Scientific, Lyon, France ⁴ Sanofi, Reading, UK	PRACTICE AREA PPD™ Evidera™ Patient-Centered Research, Thermo Fisher Scientific	CLIENT INTERNAL
--	---	---	---------------------------

Exploring the Promise of Generative Artificial Intelligence (AI) for Coding and Analysing Qualitative Patient Data: Results from a Pilot Study in Non-Hodgkin's Lymphoma

Karen Bailey,¹ Anne Skalicky,² Carla Dias Barbosa,³ Sonya Stanczyk,² Paul Cordero⁴

¹Thermo Fisher Scientific, London, UK; ²Thermo Fisher Scientific, Waltham, MA, USA; ³Thermo Fisher Scientific, Lyon, France; ⁴Sanofi, Reading, UK

Background

- Computer Assisted Qualitative Data Analysis Software (CAQDAS) programmes have integrated machine learning and artificial intelligence (AI) into their tools.
- These integrations aim to accelerate time-intensive processes related to qualitative data quality control, analysis, and reporting.
- The use of AI with sensitive patient experience data must comply with General Data Protection Regulation (GDPR) standards and documentation.

Objectives

- Evaluate how AI performs compared with human tasks and processes.
- Provide recommendations for AI use-cases that can meet regulatory standards, using both primary and secondary analysis of qualitative data.

Methods

Figure 1. Methods of primary data collection

Design	Non-interventional, cross-sectional, qualitative study.
Type of interviews	Concept elicitation (CE) and cognitive interviews (CI) to examine the relevance of PROs.
Population	US adult patients diagnosed with DLBCL or MCL.
Data collection	60-minute telephone or web-conference interviews using a semi-structured interview guide, audio recorded and transcribed verbatim, conducted between June and September 2023.
PRO being assessed	EORTC QLQ-C30, EORTC QLQ-NHL-H29, EORTC QLQ-NHL-LG29, FACT-Lym.
Objectives of the primary data collection study	<ul style="list-style-type: none">Identify relevant concepts to refine and validate existing CDMs.Evaluate selected PROs for use as endpoints in NHL clinical trials, focusing on readability, comprehensibility, relevance, and patients' expressions of items, instructions, and recall periods.
Ethics	Approval granted by US WCG IRB. Participants provided informed consent for secondary use of their data.

Abbreviations: CDM = conceptual disease model; DLBCL = diffuse large B-cell lymphoma; EORTC = European Organisation for Research and Treatment of Cancer; FACT-Lym = Functional Assessment of Cancer Therapy-Lymphoma; H29 = High Grade Module 29; IRB = Institutional Review Board; LG29 = Low Grade Module 29; MCL = mantle cell lymphoma; NHL = Non-Hodgkin lymphoma; PRO = patient-reported outcome; QC = quality control; QLQ = Quality of Life Questionnaire; WCG IRB = Western Institutional Review Board; Cognitive Group

- ATLAS.ti was the CAQDAS used for the case study (Figure 2).

The human coding and the AI coding approaches were compared in terms of the:

- Time and human resources required to finalise coding of one transcript and all the transcripts analysed
- Coding issues identified in quality control (QC)
- Number and type of codes applied to one selected transcript
- Relevance of final analysis in relation to the study objectives

Figure 2. Data analysis approaches

No. transcripts analysed	Human Analysis (ATLAS.ti v22): 30	AI Coding (v25): 30	Intentional AI Coding: 16 (due to AI restrictions)
Coding	<ul style="list-style-type: none">CE: Deductive (interview guide and CDM) and inductive (participant-driven) codes.CI: Deductive codes to ensure comprehension and relevance of the items, instrument instructions, and response options.Dual coding of first transcript.Coding monitored by a qualitative data manager.ICA of 80% was achieved for the first transcript.	<ul style="list-style-type: none">AI coding function.Human review not completed.	<ul style="list-style-type: none">Intentional AI coding function.Researchers guided by entering a summary of the interview guide questions and two instruction statements.Researchers reviewed the code dictionary and coding at a high level.One transcript underwent detailed human QC.
QC			

Abbreviations: AI = artificial intelligence; CDM = conceptual disease model; CE = concept elicitation; CI = cognitive interview; ICA = intercoder agreement; QC = quality control

About ATLAS.ti (v25) AI function

- ATLAS.ti v25 (released in April 2024) has an AI function that uses OpenAI to perform qualitative analysis.
- If the AI coding function is not enabled, no data is uploaded to OpenAI.
- Before starting AI coding, ATLAS.ti asks for your consent before uploading data to ATLAS.ti servers, as well as OpenAI servers.
- The proprietary arrangement with OpenAI ensures that any data provided is not stored and will not be used to train the OpenAI large language model (LLM).
- AI coding:** Codes inductively without human input to help categorise, interpret, and make sense of data as part of initial cycles of coding.
- Intentional AI coding:** Allows human input of "intentions," which are akin to prompts and context about the research, to inform the automatic coding.

Results

Figure 3. Time and human resources* required to finalise coding of one transcript and all the transcripts analysed

No. codes and categories	Human Coding (29 transcripts): 1,298 codes grouped into 40 categories	AI Coding (29 transcripts): 2,318 codes grouped into 578 categories	Intentional AI Coding (16 transcripts): 2,350 codes grouped into 27 categories
Time to code one transcript (4,340 words)	~103 minutes	50 seconds (no human QC)	Human QC of coding: ~180 mins
Time/human resources for initial coding	Draft code book: ~12 hours QC of coding and ICA: ~30 hours	AI function: ~10 minutes	Human development of intentions (prompts) to help generate code categories: ~30 mins AI function: ~15 minutes
Time/human resources for QC and finalise coding	QC of coding and ICA: ~30 hours	Human review was not deemed feasible due to the excessive codes generated	Human review of all transcripts to remove codes only applied to interviewer questions (Figure 5): ~10 mins Estimated time to QC all transcripts: ~48 hours

*For human coding, the resources consisted of two senior and four junior qualitative researchers. For both AI approaches, the resources consisted of two senior qualitative researchers. Abbreviations: AI = artificial intelligence; ICA = intercoder agreement; QC = quality control

Results (cont.)

Figure 4. Coding issues identified in QC of AI Methods

Coding Approach	Human Coding	AI Coding	Intentional AI Coding
Coding issues	<ul style="list-style-type: none">Amount of text coded to particular concept.Differences between codes on the interpretation of inductive concepts related to emotional wellbeing e.g., being upset, positivity.	<ul style="list-style-type: none">Application of codes to interview questions without participant responses.Application of irrelevant codesGeneral coding of symptoms "health concerns," "physical symptoms," "physical ailments"Ability to discern between similar concepts at the category or code level.	<ul style="list-style-type: none">Application of multiple wrong codes to a quotation (Figure 5, Example 1).Double coding of concept as a symptom and impact.Ability to determine which PRO the participant was describing (Figure 5, Example 2).Ability to discern between similar concepts at the category or code level (e.g., "absence of impact" and "absence of impacts" or "bother rating" and "botherlessness rating").

Abbreviations: AI = artificial intelligence; PRO = patient-reported outcome

Figure 5. Examples of issues with the Intentional AI Coding

Example 1. Application of multiple inappropriate codes to a quotation (transcript text to the left, codes applied to the right)	<ul style="list-style-type: none">DLBCL Impacts: No ImpactsMCL Impacts: Urinary IssuesEORTC QLQ-NHL-H29: No SymptomsFACT-Lym: No SymptomsFACT-Lym: No SymptomsFACT-Lym: No SymptomsDLBCL Impacts: No ImpactsDLBCL Impacts: Urinary IssuesDLBCL Impacts: Urinary IssuesDLBCL Impacts: Urinary IssuesDLBCL Impacts: Urinary IssuesDLBCL Impacts: Urinary IssuesDLBCL Impacts: Urinary Issues
Example 2. Incompletely coding statements associated with the PRO questionnaire (transcript text to the left, codes applied to the right)	<ul style="list-style-type: none">EORTC QLQ-NHL-LG29 Views: FeedbackEORTC QLQ-NHL-LG29 Views: Overall ImpressionEORTC QLQ-NHL-LG29 Views: Ease of UseEORTC QLQ-NHL-LG29 Views: Overall Impression

Abbreviations: DLBCL = diffuse large B-cell lymphoma; EORTC = European Organisation for Research and Treatment of Cancer; FACT-Lym = Functional Assessment of Cancer Therapy-Lymphoma; H29 = High Grade Module 29; LG29 = Low Grade Module 29; QLQ = Quality of Life Questionnaire; Green text indicates correct codes; Grey text indicates incorrect codes

Abbreviations: DLBCL = diffuse large B-cell lymphoma; EORTC = European Organisation for Research and Treatment of Cancer; FACT-Lym = Functional Assessment of Cancer Therapy-Lymphoma; H29 = High Grade Module 29; LG29 = Low Grade Module 29; QLQ = Quality of Life Questionnaire; Green text indicates correct codes; Grey text indicates incorrect codes

Abbreviations: DLBCL = diffuse large B-cell lymphoma; EORTC = European Organisation for Research and Treatment of Cancer; FACT-Lym = Functional Assessment of Cancer Therapy-Lymphoma; H29 = High Grade Module 29; LG29 = Low Grade Module 29; QLQ = Quality of Life Questionnaire; Green text indicates correct codes; Grey text indicates incorrect codes

Figure 6. Number and type of codes applied to one selected transcript 001-009

Coding Approach	Human Coding	Intentional AI Coding
No. of total codes applied	112	Before QC: 161; After QC: 91 (76 codes removed, 41 newly created codes, 25 already existing codes applied, 5 codes renamed)
Strengths	<ul style="list-style-type: none">Accurate coding suitable in frequency counts.Identification of symptom/impacts as spontaneous or coded.Superior application of detailed deductive codes for CI.Ability to apply latent codes (beyond description).Lacked the identification of some inductive (participant-driven) codes in comparison to AI.	<ul style="list-style-type: none">Suggestion of inductively driven relevant codes for symptoms and impacts not identified in human coding.For example, coding of "believe" in a question was also double coded as "persistent tiredness" and "inadequate rest."
Weaknesses		<ul style="list-style-type: none">Relied on substantial human input to correct AI codes (renaming/creating codes).Unable to identify the correct PRO questionnaire under discussion.Most questions required adjustment to incorporate participant response to interviewer question.

Abbreviations: AI = artificial intelligence; PRO = patient-reported outcome

Figure 7. Relevance of final analysis in relation to the study objectives

Human Coding	AI Coding
<ul style="list-style-type: none">Identification and frequency of signs and symptoms and revision of the CDMs.Mapping of participant-reported concepts with the items covered by the PROs to draw conclusion on the relevance of PROs for each population.Confirmation of content validity of the PROs.	<ul style="list-style-type: none">Recognition of relevant concepts not identified in the human coding, which can be useful for the creation of the initial code book.However, due to the issues identified with the AI coding, the provision of concept frequency (symptoms and impacts) and evidence of content validity of PROs could not be provided.

Abbreviations: AI = artificial intelligence; CDM = conceptual disease model; PRO = patient-reported outcome

Conclusions

- Given the large environmental impacts of AI, its adoption in qualitative research is justified only by the demonstration of time-saving efficiencies and quality improvements of outputs.
- The pilot results identified that the AI capabilities in ATLAS.ti v25 were not suitable for the production of regulatory-grade qualitative research involving both concept elicitation and cognitive interviews and did not create time-saving efficiencies when compared with human coding.
- The AI strengths lay in its ability to identify inductive (participant-driven) codes. Incorporating a hybrid approach combining the use of AI coding on ~2 transcripts in the development of the human-driven codebook may be useful.
- Moving forward, the ability to train the LLM underlying the AI with the outputs of the human QC would enhance the quality and efficiency of the output.
- Future work should assess how the use of AI in qualitative research complies with GDPR standards and the response of regulatory and study sponsor guidelines and policies.

Disclosures

MR, AS, CDW, and BS are employees of employees of PRO™ Evidera™ Patient-Centered Research, Thermo Fisher Scientific, and Sanofi and may hold stock or other securities in the company. This poster was funded by Thermo Fisher Scientific. Editorial and graphic design support was provided by Vanessa Carter of Thermo Fisher Scientific.

Poster Title

Evidence of Mapping Malpractice? A Review of Mapping Algorithm Usage in NICE HealthTechnology Appraisals

Objective

- To review NICE appraisals published or updated over the past 5 years to:
 - Quantify the general frequency of mapping algorithm usage
 - Highlight any trends in mapping algorithm usage across appraisals over time
 - Determine which HRQoL instruments were most frequently mapped from
 - Summarise how submitting companies identified and selected candidate algorithms for use in cost-effectiveness analyses, and highlight any relevant critiques from Evidence Assessment Groups (EAGs)

AUTHORS	AFFILIATION	PRACTICE AREA	CLIENT
George Bungey ¹ Claire Spencer ¹ Solene de Boisvilliers ² Margherita d'Errico ³ Eszter Nagy ¹	Gabor Szabo ⁴ Kerry Winter ¹ Pedro Saramago Goncalves ⁵	PPD™ Evidera™ Health Economics & Market Access, Thermo Fisher Scientific	INTERNAL

Evidence of Mapping Malpractice? A Review of Mapping Algorithm Usage in NICE Health Technology Appraisals

George Bungey,¹ Claire Spencer,¹ Solene de Boisvilliers,² Margherita d'Errico,³ Eszter Nagy,¹ Gabor Szabo,⁴ Kerry Winter,¹ Pedro Saramago Goncalves⁵
¹Thermo Fisher Scientific, London, UK; ²Thermo Fisher Scientific, Lyon, France; ³Thermo Fisher Scientific, Milan, Italy; ⁴Thermo Fisher Scientific, Budapest, Hungary; ⁵Thermo Fisher Scientific, Lisbon, Portugal

Background

- The National Institute for Health and Care Excellence (NICE) methods guidance indicates EQ-5D-3L as the preferred health-related quality of life (HRQoL) measure.¹ As EQ-5D-3L data may not always be available, mapping can be applied to bridge the evidence gap between available clinical study data and cost-effectiveness analysis requirements.
- Where EQ-5D-5L patient-level data are available, published crosswalk algorithms can be leveraged to map to EQ-5D-3L.
- Previous NICE Decision Support Unit (DSU) Technical Support Document (TSD)¹⁰ identified that mapping is used in a quarter of NICE submissions. This was published in 2011 and has since been superseded by TSD2² in 2023. While these documents^{2,3} provide guidance on conducting mapping studies, current NICE guidelines do not discuss methods for identifying and selecting mapping algorithms for use in cost-utility analyses.

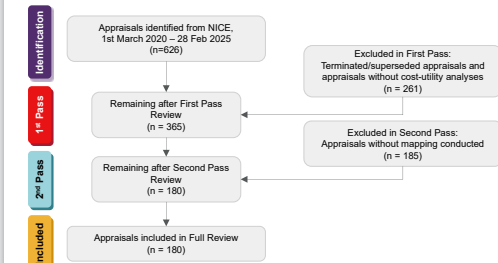
Objectives

- To review NICE appraisals published or updated over the past 5 years to:
 - Quantify the general frequency of mapping algorithm usage
 - Highlight any trends in mapping algorithm usage across appraisals over time
 - Determine which HRQoL instruments were most frequently mapped from
 - Summarise how submitting companies identified and selected candidate algorithms for use in cost-effectiveness analyses, and highlight any relevant critiques from Evidence Assessment Groups (EAGs)

Methods

- The NICE guidance website⁴ was searched to identify technology appraisals published between 01 March 2020 and 28 February 2025.
 - Terminated or superseded appraisals, as well as those without cost-utility analyses conducted, were excluded during first-pass review.
 - In the second-pass review, appraisals without mapping applied were excluded, with remaining appraisals assessed to extract information about mapping analyses conducted.
- Figure 1 presents a Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flow diagram illustrating the identification, screening, and inclusion of technology appraisals for the review.

Figure 1. PRISMA diagram



Abbreviations: NICE = National Institute for Health and Care Excellence; PRISMA = Preferred Reporting Items Systematic Reviews Meta-Analyses

Results

- Of the 626 appraisals screened, 180 (28.8%) were included for full review that reported use of mapping algorithms.
- Among the 180 appraisals included for full review, 127 (70.6%) included use of a crosswalking algorithm to map EQ-5D-5L to EQ-5D-3L.
 - 79 (62.2%) applied the van Hout et al. 2012 crosswalk algorithm.⁵
 - 47 (37.0%) used mapping algorithms from Hernandez-Alava et al. studies.⁶
 - Among these, 3 initially used van Hout et al. 2012 before switching to a Hernandez-Alava et al. study.
 - 4 (3.1%) did not clearly report which crosswalking algorithm was used in available appraisal materials.
- Figure 2 shows the proportion of the total 626 appraisals screened where mapping algorithms were used across yearly time periods:
 - Over time, no clear trend was observed in the use of mapping algorithms.
 - Following the 2022 NICE PMS36 guidance⁷ update, use of Hernandez-Alava et al. increased substantially. However, van Hout et al. 2012 was still used in 17 appraisals between March 2023 to February 2025, although this included 2 appraisals that were updated but originally published in late 2021, prior to publication of the PMS36 guidance.

Disclosures

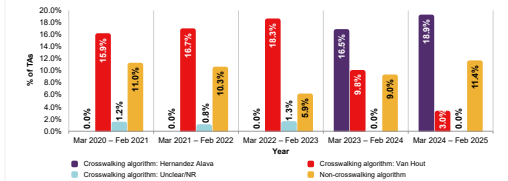
All authors are employees of PPD™ Evidera™ Health Economics and Market Access, Thermo Fisher Scientific. This poster was funded by Thermo Fisher Scientific. Editorial and graphic design support were provided by Karina Calais and Heather Nayagam of Thermo Fisher Scientific.



HTA137

Results (cont.)

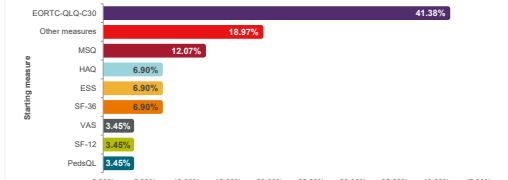
Figure 2. Percentage of TAs per year using crosswalking (Hernandez Alava, Van Hout) and non-crosswalking algorithms to map to EQ-5D-3L



Footnote: Percentages per period do not sum to 100%, as the results are derived from the total appraisals in each yearly time period. Abbreviation: TA = technology appraisal.

- 58 of the 180 appraisals included for full review (32.2%) included other mapping algorithm usage not related to EQ-5D-5L to EQ-5D-3L crosswalking:
 - Among EQ-5D-3L mappings, the most common measure was the European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire-Core 30 (EORTC-QLQ-C30) (Figure 3), mainly used in oncology appraisals (n=20) but also haematology (n=4).

Figure 3. Distribution of QoL measures mapped to EQ-5D-3L among total non-crosswalk mapping TAs



Footnote: Percentages may total above 100%, as some appraisals may mention several utility measures. Other utility measures include those used in only a single technology appraisal. Abbreviations: EORTC-QLQ-C30 = European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire – Core 30; ESS = Esworth Sleepiness Scale; HAQ = Health Assessment Questionnaire; MSQ = Migraine-Specific Questionnaire; PedsQL = Pediatric Quality of Life Inventory; QoL = quality of life; SF-12 = Short Form Health Survey 12; SF-36 = Short Form Health Survey 36; TA = technology appraisal; VAS = Visual Analogue Scale.

- 47 (81.0%) used mapped utility estimates in company base case analyses.
- 53 out of 58 (93.1%) mapped from patient-level data, although a small number (5/58; 8.6%) mapped from summary-level data; one appraisal included mapping from both.

- Submitting companies identified non-crosswalk algorithms of interest through a variety of approaches, often using previous NICE submissions for similar indications:
 - 32 (55%) used previous NICE appraisals.
 - 18 (31%) referenced identification through literature review.
 - 6 (10%) used the Health Economics Research Centre (HERC) mapping database.⁷
 - 16 (28%) did not report their approach to identification of mapping algorithms.

- A wide array of criteria for mapping algorithm selection were considered across the appraisals by submitting companies:
 - However, the application of these criteria was highly inconsistent, and only 9 out of 58 appraisals (15.5%) cited more than two selection criteria.
 - More than half (53%) either did not report selection criteria in the available documentation (21 appraisals) or relied on precedence from previous appraisals (10 appraisals).
 - Where provided, EAG feedback on mapping algorithm selection was fairly appraisal-specific rather than based on a consistent and comprehensive methodology.

Conclusions

- Our review found that mapping still plays an important role in NICE appraisals, however mapping algorithm usage not related to EQ-5D-5L to EQ-5D-3L crosswalking appeared relatively infrequent as a proportion of total appraisals.
- Although use of Hernandez Alava et al. EQ-5D-5L to EQ-5D-3L crosswalk algorithms substantially increased after the NICE PMS36 guideline update in 2022, the application of van Hout et al. 2012 was still present in some appraisals between March 2023 and February 2025.
- Algorithm identification and selection methods were inconsistent across appraisals, suggesting an unmet need for clearer guidelines.

References

- National Institute for Health and Care Excellence. NICE health technology evaluations: the manual 2022. <https://nice.org.uk/process/hmg36>.
- Lingworth L, Rowett D. NICE DSU technical support document 10: The use of mapping methods to estimate health state utility values. 2011. <https://eprints.nice.org/eprints/1444>.
- Worley A, et al. Mapping to estimate health state utility values: support document 2023. <https://eprints.nice.org/eprints/1444>.
- National Institute for Health and Care Excellence. Published Guidance, quality standards and advice. <https://nice.org.uk/quality-standards/published-guidance>.
- van Hout B, et al. Value Health. 2012; 15(1):107-16.
- University of Sheffield. Methods development. <https://sheffield.ac.uk/nice-dsu/methods-development>.
- Dakin H, et al. ERIC database of mapping studies. Version 9.0. 2023. <https://eric.nice.org.uk/eric/databases/mapping-studies>.

Poster Title

From Theory to Practice: Guidance for Identifying and Selecting Mapping Algorithms for Health Economic Models

Objective

- Conduct a targeted search to confirm the most relevant sources available for mapping algorithms and the suitability of the HERC search strategy.
- Review pertinent discussion in existing guidance on mapping algorithms to inform selection criteria as well as identification and selection approaches adopted in recent NICE appraisals.
- Leverage findings to propose guidance for identifying and selecting mapping algorithms.

AUTHORS George Bungey ¹ Caroline von Wilamowitz-Moellendorff ¹ Paulina Bajko ¹ Steven Duffy ¹ Irina Proskorovsky ²	AFFILIATION ¹ Thermo Fisher Scientific, London, UK ² Thermo Fisher Scientific, Valencia, Spain	PRACTICE AREA PPD™ Evidera™ Health Economics & Market Access, Thermo Fisher Scientific	CLIENT INTERNAL
---	---	--	---------------------------



From Theory to Practice: Guidance for Identifying and Selecting Mapping Algorithms for Health Economic Models



HTA158

George Bungey,¹ Caroline von Wilamowitz-Moellendorff,¹ Paulina Bajko,¹ Steven Duffy,¹ Irina Proskorovsky²

¹Thermo Fisher Scientific, London, UK; ²Thermo Fisher Scientific, Valencia, Spain

Introduction

The National Institute of Health and Care Excellence (NICE) methods guide recommends EQ-5D-3L as the preferred measure of health-related quality of life for health economic models.^{1,2} However, as EQ-5D-3L data on the condition, intervention, or health states of interest may not be directly available, mapping may be required to help translate data from other non-preference or preference-based measures to EQ-5D-3L (or other appropriate preference-based measures).

Despite existing NICE Decision Support Unit guidance on mapping (technical support document [TSD]¹⁰ and TSD2²), as well as other published guidance,³ there is an unmet need for guidance on identification and selection of published mapping algorithms to generate utility values for use in health economic models.

Although the Health Economics Research Centre (HERC)⁸ database provides a valuable resource for identifying mapping algorithms, the search strategy adopted has not been externally validated and may not guarantee comprehensive coverage. Supplementary literature review may also be necessary for health technology assessment (HTA) submission purposes to ensure that searches are contemporaneous.

Objectives

- Conduct a targeted search to confirm the most relevant sources available for mapping algorithms and the suitability of the HERC search strategy.
- Review pertinent discussion in existing guidance on mapping algorithms to inform selection criteria as well as identification and selection approaches adopted in recent NICE appraisals.
- Leverage findings to propose guidance for identifying and selecting mapping algorithms.

Methods

- Mapping algorithm identification:
 - A targeted review was conducted to capture publications focusing on identifying and selecting mapping algorithms.
 - The HERC database is a valuable resource for identifying mapping algorithms. We reviewed the searches it used for appropriateness and comprehensiveness and concluded that while the searches might benefit from certain sensitivity enhancements, they were sufficiently robust for identifying relevant publications.
 - The search strategy was run from January 2023 (the date of the last searches run by HERC) to September 2025. The search strategy was translated and implemented in Embase and MEDLINE via Ovid for comprehensive retrieval.
 - Keywords related to mapping, crosswalk, utility transfer, EQ-5D, or EuroQol were used to identify relevant publications, without restriction by indication or geographic region.
 - Findings from a parallel study were also incorporated, reviewing identification criteria used in NICE technology appraisals published or updated between March 2020 and February 2025.
- Mapping algorithm selection:
 - Existing published guidance on mapping (NICE TSD1¹⁰, NICE TSD2² and Petrou 2015³) was reviewed to identify any potentially relevant criteria for mapping algorithm selection.
 - Findings from a parallel study describing selection criteria used in NICE technology appraisals published or updated between March 2020 and February 2025 were also used.

Results

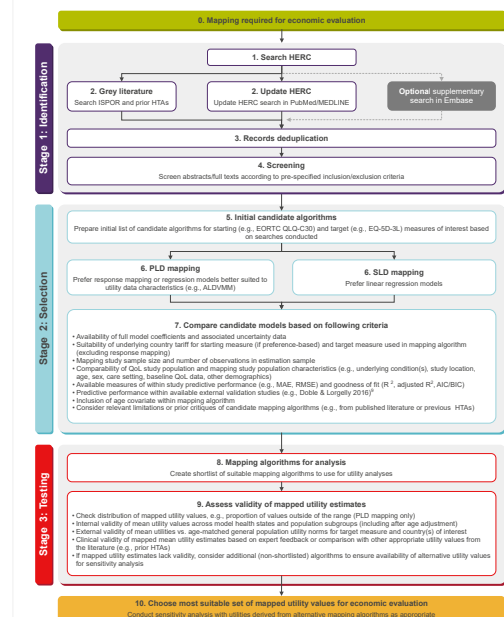
- Mapping algorithm identification:
 - It is recommended that the searches are rerun from January 2023 onwards to ensure inclusion of recent publications and that records are screened by a single researcher with at least 10% of identified publications reviewed by second screener.
 - The results indicated limited differences in identification of relevant mapping publications between PubMed and MEDLINE/Embase. However, complementary searches in MEDLINE/Embase captured additional relevant conference proceedings not indexed in PubMed.
 - A rerun of the PubMed search for the period 2023–2025 retrieved 137 records. The Medline search yielded 135 records, closely aligning with PubMed results, as expected. The Embase search retrieved 276 records. After deduplication, the combined total across all sources was 286 unique records.
 - A review of these additional publications identified ten studies containing mapping algorithms with model coefficients provided: four full publications and six International Society for Pharmacoeconomics and Outcomes Research (ISPOR) abstracts/posters.
 - None of the full publications were retrieved by the original HERC search, despite being indexed in PubMed, suggesting some potential limitations in the original search strategy.
 - While Medline and Embase both captured relevant mapping studies, Embase captured numerous additional conference abstracts and posters, as reported above.
 - Results of the parallel study reviewing recent NICE appraisals highlighted that although the HERC database, literature reviews, prior NICE appraisals, and external validation studies were considered as sources to identify mapping algorithms, they were inconsistently used across appraisals.²
- Mapping algorithm selection:
 - Review of the published guidance highlighted the need to consider the type of mapping model used, measures of predictive performance/model fit, study population characteristics, and the potential value of age as a covariate within mapping models.
 - Excluding EQ-5D-5L to EQ-5D-3L cross-walk algorithm usage, results of the parallel study reviewing NICE appraisals between March 2020 and February 2025 indicated that selection criteria were inconsistently applied with more than half either not reporting selection criteria or relying on precedents from previous NICE appraisals.
 - Although NICE DSU TSD1¹⁰ and TSD2² outline important limitations of linear mapping algorithms, as highlighted in NICE technology appraisal (TA)840,⁹ there may be instances where patient level quality of life (QoL) data are not available from the trial nor appropriate utilities from the literature, and where mapping of summary level non-preference based QoL data from the literature may be more suitable. Due to the non-collapsible nature of non-linear models, linear models may be preferred in such cases.

Results (cont.)

Assessing the validity of the mapped utility estimates is valuable to help determine the most suitable mapping algorithm. However, exploration of the full range of initial candidate models may not be practical for measures where a large number of mapping algorithms are available (e.g., QoL-C30). Thus, it may be more pragmatic to conduct an initial selection phase to shortlist mapping algorithms before validating mapped utility estimates. If the shortlisted algorithms fail to produce suitable utility estimates, the excluded algorithms can be reconsidered.

Informed by the findings outlined above, a set of pilot guidance was then developed outlining a practical process for identification, selection, and testing of mapping algorithms (Figure 1).

Figure 1. Mapping Algorithm Identification and Selection Flow Diagram



Abbreviations: AIC = Akaike information criterion; ALD/VM = adjusted limited dependent variable measure models; BIC = Bayesian information criterion; EORTC QLQ-C30 = European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire; HTA = health technology assessment; HERC = Health Economics Research Centre; ISPOR = International Society for Pharmacoeconomics and Outcomes Research; MAE = mean absolute error; PLD = patient level data; QoL = quality of life; RMSE = root mean squared error; SLD = summary level data

Conclusions

- Clear reporting of mapping methods used to inform utility analyses for economic models is essential to ensure transparency and reproducibility.
- The proposed pilot guidance aims to offer a clear, comprehensive process for mapping algorithm identification and selection.

References

1. National Institute for Health and Care Excellence. NICE health technology evaluations: the manual (PMG36) 2022. <https://nice.org.uk/procurement/36/>
2. National Institute for Health and Care Excellence. Guide to the methods of technology appraisal 2013 (PMG59) 2013. <https://nice.org.uk/guidance/9781471300000>
3. Longford I, Hays RD. Technical support document: The use of mapping methods to estimate health state utility values. 2011. <https://www.nhs.uk/guidance/NG142544/technical-support-document-2011-01-14>
4. Vassiliou A, et al. Mapping to estimate health state utility values: technical support document 27. <https://www.nhs.uk/guidance/NG142544/technical-support-document-2011-01-14>
5. Petrou S, et al. Pharmacoeconomics. Oct 2015;33(10):993-1011. doi: 10.1007/s11111-015-0411-1
6. Dainoff H, et al. ERIC database of mapping studies, Version 9.0 2023. <https://www.herc.co.uk/uknowledgehealtherec-database-of-mapping-studies-2023-09-01>
7. International Society for Pharmacoeconomics and Outcomes Research. ISPOR Europe 2025. November 2025. Glasgow, Scotland. National Institute for Health and Care Excellence. <https://www.nice.org.uk/guidance/TA840>
8. Dainoff H, Longford I. *Qual Life Res*. 2015;23(4):951-971.

Disclosures

All authors are employees of PPD™ Evidera™ Health Economics & Market Access, Thermo Fisher Scientific. This poster was funded by Thermo Fisher Scientific.

Acknowledgments

Editorial and graphic design support were provided by Karissa Calera, Linda Allen, and Kaitlyn Nakayama of Thermo Fisher Scientific.



Poster Title

Anchors Away: Navigating Unanchored Indirect Comparisons with Multilevel Unanchored Meta-regression

Objective

- To address this gap, we introduce multilevel unanchored meta-regression (ML-UMR)—a novel extension of ML-NMR for unanchored comparisons—and assess its performance via simulation.

AUTHORS Conor Chandler ¹ K. Jack Ishak ²	AFFILIATION ¹ Thermo Fisher Scientific, Waltham, MA, USA ² Thermo Fisher Scientific, Montreal, Canada	PRACTICE AREA PPD™ Evidera™ Health Economics & Market Access, Thermo Fisher Scientific	CLIENT INTERNAL
---	--	--	---------------------------

Anchors Away: Navigating Unanchored Indirect Comparisons with Multilevel Unanchored Meta-regression

Conor Chandler,¹ K. Jack Ishak²
¹Thermo Fisher Scientific, Waltham, MA, USA; ²Thermo Fisher Scientific, Montreal, Canada

Background

- Imbalances in clinically relevant baseline characteristics (i.e., treatment effect modifiers and/or prognostic factors [PFs]) across study populations may lead to biased results in unadjusted indirect comparisons (e.g., network meta-analysis).
- Multilevel network meta-regression (ML-NMR) enhances network meta-analysis by statistically adjusting for effect modification (EM) in connected networks.
- ML-NMR is the most flexible anchored population-adjusted indirect comparison (PAIC) method, as it enables transporting effect estimates to any target population of interest and is capable of comparing any number of treatments.
- Existing PAIC methods for disconnected networks (i.e., unanchored matching-adjusted indirect comparison [MAIC] and simulated treatment comparison [STC]), however, are limited to pairwise comparisons and cannot transport estimates beyond the comparator population.

Objectives

- To address this gap, we introduce multilevel unanchored meta-regression (ML-UMR)—a novel extension of ML-NMR for unanchored comparisons—and assess its performance via simulation.

Methods

ML-UMR: A Novel Extension of ML-NMR

- We introduced two types of Bayesian ML-UMR models (Figure 1) and motivated the application with simulated examples of pairwise PAICs for treatment A (index) vs. B (comparator). Here, patient-level data (PLD) are available for A, while only aggregate-level data (ALD) are reported for B.
- The first type of ML-UMR model, invoking the shared PF assumption (SPFA), implied there is treatment effect homogeneity at the individual level (i.e., no EM) for A vs. B.
 - Fitted using PLD for A and ALD for the overall population for B.
- The second type of ML-UMR model, relaxing SPFA, allowed for potential treatment effect heterogeneity at the individual level (i.e., EM) for A vs. B.
 - Fitted using PLD for A and ALD for non-overlapping subgroups that partition the overall population for B (e.g., four subgroups formed by the interactions between PFs X_1 and X_2 in simulated examples).
- The general formulas for an ML-UMR indirectly comparing A vs. B are presented in Figure 1, which extend straightforwardly to analyses involving additional treatments and/or multiple studies per treatment.

Figure 1. ML-UMR Model Comparing A vs. B

	ML-UMR Type 1 (Invoking SPFA)	ML-UMR Type 2 (Relaxing SPFA)
Individual-level component	$g(\theta_{i(A)}) = \alpha_{i(A)} + x_{i(A)}^T \beta_{i(A)}$	$g(\theta_{i(A)}) = \alpha_{i(A)} + x_{i(A)}^T \beta_{i(A)}$
Aggregate-level component	$\theta_{i(B)} = \int g^{-1}(\alpha_{i(B)} + x_{i(B)}^T \beta_{i(B)}) f(x_{i(B)}) dx$ $\theta_{i(B)} = \sum_{j=1}^J g^{-1}(\alpha_{i(B)} + x_{i(B)}^T \beta_{i(B)}) \left(\frac{1}{n_{i(B)j}} \right)$	$\theta_{i(B)} = \int g^{-1}(\alpha_{i(B)} + x_{i(B)}^T \beta_{i(B)}) f(x_{i(B)}) dx$ $\theta_{i(B)} = \sum_{j=1}^J g^{-1}(\alpha_{i(B)} + x_{i(B)}^T \beta_{i(B)}) \left(\frac{1}{n_{i(B)j}} \right)$

Abbreviations: ML-UMR = multilevel unanchored meta-regression; SPFA = shared prognostic factor assumption. $\alpha_{i(A)}$ and $\alpha_{i(B)}$ are individual- and aggregate-level distributions in $g(\cdot)$ (normal and Binomial in simulation), β represents the mean outcome, x is a vector of PFs, $\beta_{i(A)}$ is a vector of regression coefficients for PFs, α is the baseline outcome for treatments A and B , i is the number of individuals in the index study (A), j is an indicator for the mutually exclusive subgroups for each possible combination of PFs for B . Note the individual-level and aggregate-level components of ML-UMR presented here are for a single pairwise PAIC of A vs. B. These formulas can be extended and written more generally to handle three or more treatments and multiple studies per treatment.

Simulation Study

- Our simulation study indirectly compared binary outcomes for A vs. B from two single-arm studies.
- Population imbalance was induced across studies by generating correlated PFs with different means, and outcomes were simulated: 1) assuming PFs have the same effect on outcomes for A and B (i.e., SPFA), which implies no EM, and 2) relaxing SPFA, thereby inducing weak to strong EM.
- Bayesian ML-UMR models were fitted using OpenBUGS (3 chains, 2,000 total iterations per chain, 1,000 burn-in period) to assess the absolute and relative bias and coverage of 95% credible intervals (CrIs) of predicted marginal log odds ratios (LORs) in the comparator and index populations.
- The two types of ML-UMR models were evaluated under three scenarios: 1) shared PFs for A and B (i.e., no EM); 2) weak violation of SPFA (i.e., weak EM); and 3) strong violation of SPFA (i.e., strong EM).
- Table 1 summarizes the assumptions and simulation settings for this study.

Table 1. Assumptions for Simulation Study

Parameter	Assumption
Monte Carlo replications	500
Sample sizes	Equal samples of $n=1,000$ for each trial
PFs	Two binary variables (X_1, X_2) with moderate correlation (0.5) Index: $X_1 = X_2 = 0$ Comparator: $X_1 = X_2 = 0.7$
Prognostic strength of X_1 *	Scenarios 1-3: $\beta_{1,A} = \beta_{1,B} = -1$
Prognostic strength of X_2 *	Scenario 1: $\beta_{2,A} = \beta_{2,B} = -2$ Scenario 2: $\beta_{2,A} = -2, \beta_{2,B} = -1.75$ Scenario 3: $\beta_{2,A} = -2, \beta_{2,B} = -1$
Baseline outcome*	$\alpha_A = 1, \alpha_B = 0.25$

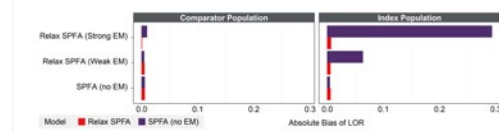
Abbreviation: PF = prognostic factor
*Modelled on the logit scale

Results

Simulation Study

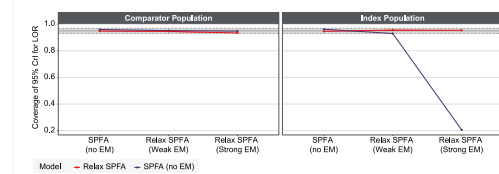
- ML-UMR models invoking SPFA accurately predicted LORs for A vs. B in the comparator population regardless of the true EM strength (bias<0.009 [-6.4%] and 94.8%–96.0% coverage) (Figure 2 and Figure 3).
- The predicted LORs for A vs. B in the index population were unbiased in the absence of EM and relatively robust to weak EM (bias=-0.06 [-1.1%]; coverage=93%); a high degree of bias (-0.29 [-67%]) was observed, however, when EM was strong.
- Relaxing SPFA in the ML-UMR model resulted in accurate LORs in both the index and comparator populations across all scenarios: bias<0.007 (<1.5%) and 93.4%–95.6% coverage.

Figure 2. Absolute Bias of LOR



Abbreviations: EM = effect modification; LOR = log odds ratio; SPFA = shared prognostic factor assumption. The absolute bias represents the difference between the average prediction of the ML-UMR model and the true treatment effect (ignoring the direction of error), where values closer to 0 indicate that the average model predictions more closely reflect reality and large values indicate a systematic error with model predictions (i.e., over- or under-prediction of the true effect). Note, the predicted effect in the index is biased when SPFA is incorrectly assumed in the model, but the estimate in the comparator population has negligible bias. This is because the regression coefficients for PFs are more heavily influenced by the PLD for A, and do not generalize to B outside of the comparator population.

Figure 3. Coverage of 95% CrIs for LOR



Abbreviations: CrI = credible interval; EM = effect modification; LOR = log odds ratio; SPFA = shared prognostic factor assumption. Empirical coverage probabilities of estimated 95% CrIs for LORs. Estimates within the shaded region did not significantly differ from the nominal confidence level of 95%. Note, the assumed sample sizes were large in the simulation study ($n=1,000$ for each arm) and other simplifying assumptions were imposed, which may contribute to the observed coverage probabilities.

Conclusions

- This study demonstrates that the ML-NMR framework can be extended for unanchored indirect comparisons.
- In the pairwise unanchored setting, the SPFA is the fundamental assumption for transporting effect estimates from the comparator population to a different target population, such as the index.
 - This is analogous to the shared effect modifier assumption (SEMA) in anchored PAICs. SPFA is a stronger assumption than SEMA, and thus it may be more difficult to achieve transportability in unanchored analyses.
 - The SPFA may be relaxed by leveraging comparator data for appropriate subgroups and/or multiple comparator studies² or clinical expert opinion. Stratified analyses could also be explored. Further research is required in this area.
 - Obtaining the data necessary to relax SPFA will be challenging in practice, as the reporting of comparator data is often limited.
- ML-UMR effectively simplifies to unanchored STC if two treatments are being compared in the comparator population.
 - Consistent with the findings of Ren et al.,² effect estimates in the comparator population were unbiased if all PFs were included in the model, regardless if the SPFA is violated.
- Unlike MAIC and STC, ML-UMR allows transporting relative effect estimates to any target population under certain assumptions (e.g., the SPFA may be required if sufficient data are not reported for comparators), can compare any number of treatments, and can synthesize multiple studies per treatment, but it remains subject to strong limitations inherent to unanchored comparisons.

References

- Phillips DM, et al. *J R Stat Soc Ser A Stat Soc*. 2020; Jan;183(3):1189-1210. 2. Phillips, DM. PhD thesis, University of Bristol. 2019.
- Ren S, et al. *Res Synth Methods*. 2024; Jul;15(4):657-670.

Disclosures

All authors are employees of PPD™ Evidera™ Health Economics & Market Access, Thermo Fisher Scientific. This poster was funded by Thermo Fisher Scientific.

Acknowledgments

Editorial and graphic design support were provided by Sean Smith and Dawn Gargner of Thermo Fisher Scientific.

Poster Title

An Exploration of Implementing Herd Immunity Into Static Models for Economic Evaluation of Vaccines


Objective

- Aim: Compare vaccination effects on infection rate in both DTMs and Markov models.
- Develop and compare two models for the same hypothetical disease:
 - DTM: Captures person-to-person transmission and indirect (herd) effects through changing prevalence and contacts
 - Markov model: Models infection using a simpler time-step (discrete time), with transition probability from susceptible to infected informed by the DTM critiques from Evidence Assessment Groups (EAGs)
- Evaluate impact of herd immunity on infection rate
- Analyze how different functional forms of herd effect in the Markov model influence results

<p>AUTHORS Josie Dodd Des Dillon-Murphy Ruth Chapman</p>	<p>AFFILIATION Thermo Fisher Scientific, London, UK</p>	<p>PRACTICE AREA PPD™ Evidera™ Health Economics & Market Access, Thermo Fisher Scientific</p>	<p>CLIENT INTERNAL</p>
--	---	---	--

An Exploration of Implementing Herd Immunity Into Static Models for Economic Evaluation of Vaccines

Josie Dodd,¹ Des Dillon-Murphy,¹ Ruth Chapman¹
¹Thermo Fisher Scientific, London, UK



PT30

Background

- Infectious diseases like pneumococcal, influenza, and COVID-19 can cause serious complications or death, especially in the elderly and high-risk groups.^{1,2}
- Vaccination is highly effective in reducing clinical and economic burdens.¹
- Dynamic transmission models (DTMs) are used in infectious disease modeling but are complex and computationally intensive.
- Markov models are more common in health economic assessments but struggle with incorporating herd effects, potentially underestimating vaccine impact.
- Indirect effects in Markov models are included through fixed-input parameters from empirical data to adjust incidence rates at specific vaccination coverage levels.

Objectives

- Aim: Compare vaccination effects on infection rate in both DTMs and Markov models.
- Develop and compare two models for the same hypothetical disease:
 - DTM: Captures person-to-person transmission and indirect (herd) effects through changing prevalence and contacts
 - Markov model: Models infection using a simpler time-step (discrete time), with transition probability from susceptible to infected informed by the DTM
- Evaluate impact of herd immunity on infection rate
- Analyze how different functional forms of herd effect in the Markov model influence results

Methods

- Both models use a simple susceptible-infected-susceptible (SIS) structure (infection does not increase mortality, and recovered individuals return to the susceptible pool) illustrated in Figure 1. The population is stratified into 3 age equally sized groups: Group 1, Group 2, and Group 3.

Results (cont.)

Indirect effect

- By comparing the number of infections in the Markov and DTM, we can gauge the extent of herd immunity predicted by the DTM.
- Examining various roll-out rates helps understand their impact on herd immunity.
- Markov models often include herd immunity, approximated using real-world evidence (RWE) as a constant modifier to the incidence rate. This study used a constant value of 6.8%, based on the indirect effect of the pneumococcal conjugate vaccine.¹
- Figure 3 shows that the Markov model with constant herd immunity predicted fewer infections than the DTM before reaching 40% vaccination coverage, suggesting it was less conservative and may have overestimated herd effect. Later, the DTM predicted fewer infections, indicating the Markov model underestimated indirect effects in the later phase of the simulation. A fourth simulation added a linear vaccine-coverage-dependent herd effect to the Markov model. The results were more conservative but did not match the DTM, indicating that the herd effect was not linear, though it improved over the constant herd effect.

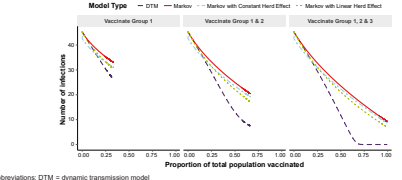
Figure 3. Incident infections for increasing Group 1 vaccination coverage for increasing rates of vaccination



Abbreviation: DTM = dynamic transmission model

- In Figure 4, we present the results of a similar analysis, where instead of varying roll-out rates, we expanded coverage to different demographic groups. When expanding vaccination coverage, we observed a more pronounced herd effect, with infections dropping to zero when less than 75% of the total population was covered. Similar to Figure 3, the Markov model with constant herd effect underestimated infections at low coverage levels and overestimated them at high coverage levels.

Figure 4. Incidence of infections for increasing population coverage under various vaccination strategies targeting different population groups



Abbreviations: DTM = dynamic transmission model

Figure 1. Model structure for both the DTM and Markov model



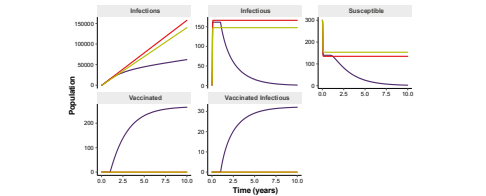
Modelling assumptions

- Demography is not included in the model.
- The rate at which patients transition from the susceptible health state to the infected health state was informed by the incidence rate in the DTM to minimize the differences between the Markov model and the DTM.
- The model was run over a period of 10 years, with vaccination commencing after a 1-year burn-in phase.
- The force of infection was assumed to be the same for all age groups, and the contact matrix used in the model was symmetric.

Results

- Figure 2: Shows the disease dynamics predicted by the DTM.
 - Before vaccination: Infectious population quickly reaches a steady state
 - Upon vaccination: Significant drop in infections in Group 1, reducing infections in Groups 2 and 3, especially in Group 2 due to higher contact rates
 - Highlights: Indirect benefits of vaccinating Group 1 in reducing overall disease burden
 - System stabilizes to a new steady state within 10 years
- Markov model simulations (not shown):
 - Demonstrates similar behavior to Figure 2
 - Notable differences: Group 2 and Group 3 are not affected by the vaccination of Group 1
 - Indicates: Indirect effects of vaccinating Group 1, inherently included in DTM, are not captured in the Markov model

Figure 2. Disease dynamics of the dynamic transmission model



Disclosures


JD, DDM and RC are employees of PPD™ Evidera™ Health Economics & Market Access, Thermo Fisher Scientific. This poster was funded by Thermo Fisher Scientific. Editorial and graphic design support were provided by Corinne Cole and Katherine Nakayama of Thermo Fisher Scientific.

References

1. Van Vliet et al. *BMC Infect Dis*. 2015;15:25. 2. Hens et al. *BMC Med Res Methodol*. 2018;18(1):62. 3. Shi T, et al. *Lancet Glob Health*. 2017;15(1):e4. 4. Ballester C. Economic Assessment of Routine PCV10 for Children. 2023. <https://doi.org/10.26434/chemrxiv-2023-06-21-12360>. Pneumococcal-Stockpile-508.pdf.

Presented at ISPOR – The Professional Society for Health Economics and Outcomes Research—Europe | 9–12 November 2025 | Glasgow, Scotland, UK

© 2025 Thermo Fisher Scientific Inc. All rights reserved. All trademarks are the property of Thermo Fisher Scientific and its subsidiaries unless otherwise indicated. This information is not intended to encourage use of these products by any receiver the right belongs to the intellectual property rights of others.





Poster Title

Using a Large Language Model (LLM) for Data Extraction of Studies: Learnings from a Targeted Literature Review (TLR) in Non-small Cell Lung Cancer (NSCLC)

Objective

- This study aimed to evaluate a zero-shot GPT-4-assisted extraction approach for clinical outcomes from RWE studies in a specific cancer subpopulation to address the following objectives:
 - Identify and document challenges during data extraction, especially in RWE designs and studies with mixed populations, focusing on subsets of interest.
 - Provide recommendations for using LLMs in these instances.

<p>AUTHORS</p> <p>Mariana Farraia¹ Anuja Pandey² Eugenia Priedane³</p>	<p>AFFILIATION</p> <p>Allie Cichewicz⁴ Caroline von Wilamowitz-Moellendorff²</p>	<p>PRACTICE AREA</p> <p>PPD™ Evidera™ Health Economics & Market Access, Thermo Fisher Scientific</p>	<p>CLIENT</p> <p>INTERNAL</p>
--	---	---	--------------------------------------


SA101

Using a Large Language Model (LLM) for Data Extraction of Studies: Learnings from a Targeted Literature Review (TLR) in Non-small Cell Lung Cancer (NSCLC)

Mariana Farraia,¹ Anuja Pandey,² Eugenia Priedane,³ Allie Cichewicz,⁴ Caroline von Wilamowitz-Moellendorff²

¹Thermo Fisher Scientific, Ede, Netherlands; ²Thermo Fisher Scientific, London, UK; ³BeOne Medicines (UK), Ltd., London, UK; ⁴Thermo Fisher Scientific, Waltham, MA, USA

Background

- Large Language Models (LLMs) streamline literature reviews by reliably extracting relevant information, thus reducing manual effort.
- We previously reported:
 - High accuracy (84%, range: 66%–96%) using zero-shot prompts for extraction of disease-specific clinical outcomes from a systematic literature review (SLR) of randomized controlled trials (RCTs) in atopic dermatitis.¹
 - High reproducibility rates (80.6%–100%) under the same human operator and conditions, but lower reliability (65.7%–95.5%) between different operators using identical prompts.²
- While extraction has been tested on a broad range of therapeutic areas using RCT data,^{3,4} real-world evidence (RWE) studies and complex research questions, such as those involving specific subpopulations, are not fully understood.

Objectives

- This study aimed to evaluate a zero-shot GPT-4-assisted extraction approach for clinical outcomes from RWE studies in a specific cancer subpopulation to address the following objectives:
 - Identify and document challenges during data extraction, especially in RWE designs and studies with mixed populations, focusing on subsets of interest.
 - Provide recommendations for using LLMs in these instances.

Methods

- A targeted literature review (TLR) assessed the comparative effectiveness and safety of immunotherapy treatments for non-small cell lung cancer (NSCLC) patients with programmed death-ligand 1 (PD-L1) expression ≥50% in RWE studies. Sixteen publications covering ten observational primary studies were included.
- Zero-shot prompts were developed, tested, and optimized for a proprietary GPT-4-based LLM using one included publication. Once satisfactory output was achieved (i.e., all predefined fields extracted with correct formatting and no critical omissions or errors after human validation), data were then extracted by the LLM for each remaining publication individually.
- Zero-shot prompt: input given to a LLM to perform a task without any prior specific training or examples for that task; the model relies solely on the prompt and the provided documents
- Data extracted by the LLM were validated by a human investigator, with the main challenges noted during the process.
- Related publications were identified and categorized manually by a human reviewer.

Results

- The main challenges identified during the extraction and validation process were:
 - Difficulties in isolating data for the target population (PD-L1 ≥50%) when reported as a subset of a broader NSCLC population with diverse PD-L1 expression statuses.
 - Incorrect or missing data extracted by the LLM for subgroups (e.g., sex, age, smoking status) and fields related to study characteristics (e.g., follow-up duration, data source).
 - Multiple files for a single study caused discrepancies from data presentation (tables, text, figures), multiple data types and timepoints, variations between main text and supplemental materials, and related publications.
- As a result of these challenges, additional extraction and re-validation of subgroup data, along with correction of formatting issues, resulted in time expenditure equal to or greater than validating manual extractions.

Limitations and recommendations

- For each identified challenge, the limitations with the extraction approach are presented in **Table 1**, along with a suggested workflow and considerations for future projects using LLMs for data extraction (**Figure 1**).

Data extraction challenge	Limitations	Recommendations
Mixed population studies: difficulties identifying data from the target population (PD-L1 ≥50%) within a broader population (NSCLC)	<ul style="list-style-type: none"> LLM often extracts data for the overall population instead of the subpopulation of interest. 	<ul style="list-style-type: none"> Develop specific prompts for studies with subpopulations of interest: <ul style="list-style-type: none"> Include detailed context and clearly indicate that only data for the target subpopulation are of interest.
Incorrect or missing data for subgroups	<ul style="list-style-type: none"> LLM may extract incorrect/incomplete subgroup data. The LLM used for this study did not recognize data in figures. 	<ul style="list-style-type: none"> Enhance prompt specificity for subgroup data. Include detailed instructions and examples in the prompt.
Lack of recognition of related publications	<ul style="list-style-type: none"> The prompt was not designed to enable the LLM to identify related publications, resulting in some redundant extractions. 	<ul style="list-style-type: none"> Consider two approaches: <ol style="list-style-type: none"> Identify related publications manually and design specific prompts to extract only new data. Identify related publications, compare data manually, and use LLM for extraction only if new data points are present, refining the prompt as needed.
Supplementary materials	<ul style="list-style-type: none"> Multiple documents (such as the main publication and its supplement) complicate LLM processing, especially in mixed population studies with extensive subgroup data. 	<ul style="list-style-type: none"> Enhance prompt instructions for handling multiple documents. Specify prioritized information for extraction. Note relevant subpopulation data may be in supplementary material.
Formatting issues	<ul style="list-style-type: none"> Inconsistent formatting in extracted data requires additional processing and standardization. 	<ul style="list-style-type: none"> Standardize format requirements in prompts. Provide clear examples.
Generic platform vs platform integrated within literature review software	<ul style="list-style-type: none"> Using a generic LLM interface may increase time requirements due to manual steps needed to use it for literature reviews (e.g., identifying related publications, providing examples in prompts). 	<ul style="list-style-type: none"> LLMs integrated into a literature-review platform are trained on structured research data extraction variables and outcomes with higher fidelity for TLRs. Built-in linkage across related publications reduces redundant data extraction from overlapping or duplicate studies.

Note: Limitations may not be generalizable across various LLM architectures, implementations, or platform integrations. Abbreviations: LLM = large language model; NSCLC = non-small cell lung cancer; PD-L1 = programmed death-ligand 1

Figure 1. Suggested workflow diagram and considerations for future projects using LLM for data extraction



Some LLMs now support reading Excel™ files. If available, uploading the Excel extraction table for direct population may streamline the workflow. At the time of this work, this functionality was not available. Abbreviations: LLM = large language model.

Conclusions

- LLMs can extract data from RWE studies but face significant challenges with clinical outcomes, especially for subgroups and mixed populations. Heterogeneity and limited standardization in reporting of observational studies likely contributed to errors in LLM-assisted extraction with zero-shot prompts.
- Further research into prompt engineering (i.e., design and execution) is needed to improve efficiency and accuracy in LLM-assisted data extraction for RWE study designs and complex research questions, particularly across multiple files per study and various types of data. Additional considerations are needed for generic LLMs versus SLR-specific tools.
- Future projects should consider and evaluate the time required for human re-extraction, re-validation, and prompt development/optimization to accurately assess potential time savings.

References

- Shree A, et al. Value Health. 2024;27(12):1547S.
- Shree A, et al. Value Health. 2023;26(9):1209S.
- Gartshofer G, et al. Res Synth Methods. 2024;15(4):576-89.
- Schmitt L, et al. PLOSOne. 2021;16:1-11.

Disclosures

Mr. An. CvMM is an employee of PPD™ Evidera™ Health Economics & Market Access, Thermo Fisher Scientific. AC was employed by Thermo Fisher Scientific at the time this study was conducted. This poster was funded by Thermo Fisher Scientific.



Presented at ISPOR —The Professional Society for Health Economics and Outcomes Research—Europe | 9–12 November 2025 | Glasgow, Scotland, UK





Poster Title

Real-world Evidence on the Broader Benefits of GLP-1 RA-based Therapies: A Global Targeted Literature Review

Objective

- The goal of the study was to review published real-world benefits of GLP-1 RA-based therapies beyond their established effects on weight loss and glycaemic control.

<p>AUTHORS</p> <p>Katherine Gibbs Naomi Stapleton Magdalena Monica Roy Mootoosamy Eoin Jennings</p>	<p>AFFILIATION</p> <p>Thermo Fisher Scientific, London, UK</p>	<p>PRACTICE AREA</p> <p>PPD™ Evidera™ Health Economics & Market Access, Thermo Fisher Scientific</p>	<p>CLIENT</p> <p>INTERNAL</p>
--	---	---	--------------------------------------



Real-world Evidence on the Broader Benefits of GLP-1 RA-based Therapies: A Global Targeted Literature Review

CO206

Katherine Gibbs,¹ Magdalena Monica,¹ Eoin Jennings,¹ Naomi Stapleton,¹ Roy Mootoosamy¹
¹Thermo Fisher Scientific, London, UK

Background

- The glucagon-like peptide-1 receptor agonist (GLP-1 RA) semaglutide (Wegovy[®]) and dual GLP-1 RA/glucocorticoid-dependent insulinotropic polypeptide (GIP) agonist tirzepatide (Mounjaro[®]) have transformed the treatment of overweight/obesity.^{1,2}
- However, health technology assessment (HTA) bodies have criticised the evidence for longer-term impacts on weight-related comorbidities as insufficient.^{3,4}
- As weight-related comorbidities are a major driver of the economic burden of overweight/obesity,^{5,6} understanding the broader impact of GLP-1 RA-based therapies may better capture their value to healthcare systems and society.

Objectives

- The goal of the study was to review published real-world benefits of GLP-1 RA-based therapies beyond their established effects on weight loss and glycaemic control.

Methods

- A targeted literature review was conducted in Embase and Medline using a predefined search strategy for real-world evidence (RWE) on semaglutide or tirzepatide published through 4 June 2025.
- No restrictions were applied to the patient population.
- Records reporting clinical outcomes beyond body weight/composition and glycaemic control were included.
- For each study, data from the longest observation period per endpoint were analysed qualitatively.
- Results are presented as the number of studies showing improvement among those reporting a given endpoint and comparison (n/N).
- Detailed results per included study can be obtained by scanning the quick response (QR) code below.

Results

- Of 580 records screened, 186 advanced to full-text screening and 133 were included. The number of unique studies identified for each outcome and the number of reported endpoints are summarised in **Figure 1**.

Cardiovascular Outcomes

- GLP-1 RAs improved systolic, diastolic and pulmonary artery pressure (each 1/1) vs. baseline, and arrhythmia (2/2), heart failure (HF), HF exacerbation, atrial fibrillation, ischaemic heart disease, peripheral vascular disease, cardiovascular (CV) disease, myocardial infarction, major adverse cardiovascular events (MACE) and congestive heart disease (each 1/1) vs. no GLP-1 RA, but showed no effect on ischaemic stroke (0/1).
- Semaglutide improved N-terminal pro-hormone of brain natriuretic peptide (NT-proBNP) (2/2), New York Heart Association (NYHA) class (2/2), 6 Minute Walk Test (6MWT), atherogenic index of plasma, carotid intima-media thickness, HF-related emergency department (ED) visits, HF-related hospitalisation, Kansas City Cardiomyopathy Questionnaire (KCCQ), lipid triad index, risk of atherosclerotic CV disease, uric acid/high-density lipoprotein (HDL) ratio (each 1/1), systolic blood pressure (SBP) (2/2) and diastolic blood pressure (DBP) (1/1) vs. baseline, and improved SBP (2/2) and reduced the risk of ischaemic stroke (3/4) vs. no GLP-1 RA.
- Tirzepatide improved 6MWT, KCCQ, left ventricular ejection fraction, NT-proBNP, NYHA class, heart rate (each 1/1), SBP (5/5) and DBP (3/4) vs. baseline, and improved HF exacerbation (1/1), MACE (2/2) and a composite endpoint of HF exacerbation and all-cause mortality (ACM) (1/1), stabilised SBP (1/1) and DBP (1/1) and reduced the risk of ischaemic stroke (1/1) vs. no GLP-1 RA-based therapy.
- Tirzepatide reduced the incidence of cerebral infarction vs. semaglutide (1/1) but there was no difference in ischaemic stroke, acute coronary syndrome, HF or ischaemic heart disease (each 0/1).

Renal Outcomes

- GLP-1 RAs improved blood urea nitrogen (1/1) and renal resistive index (1/1) vs. baseline. No improvements were reported for albuminuria (0/1), creatinine (0/4), urine albumin-creatinine ratio (UACR) (0/1) or urea (0/1). Improvements were observed in acute kidney injury (1/1) and renal resistive index (1/1), with mixed results for estimated glomerular filtration rate (eGFR) (1/4), vs. no GLP-1 RAs.
- Semaglutide improved UACR (13/19) and uric acid (5/10) vs. baseline. Few studies reported improvements in chronic kidney disease (CKD) status (1/1), creatinine (3/18), albuminuria (0/1), blood urea nitrogen (0/2), microalbuminuria (0/1) or eGFR (1/2). Improvements were reported in the risk of acute kidney injury (2/3), with mixed results for eGFR (1/2), vs. no GLP-1 RA.
- Tirzepatide improved eGFR (3/4) but not creatinine (0/1) vs. baseline. Improvements were noted in eGFR (1/1) and major adverse kidney events (1/1) vs. no GLP-1 RA-based therapy.
- Tirzepatide showed improvements vs. GLP-1 RAs in the risk of acute kidney injury (2/2), kidney events (1/1), major adverse kidney events (1/1) and UACR (1/1).

Hepatic Outcomes

- GLP-1 RAs had mixed effects on alkaline phosphatase (ALP) (1/2), alanine transaminase (ALT) (2/6) and ultrasound attenuation parameter (1/1) vs. baseline, and no impact on aspartate aminotransferase (AST) (0/5), enhanced liver fibrosis score (0/1), gamma-glutamyl transferase (GGT) (0/1), liver stiffness (0/1) or total bilirubin count (0/2).
- Semaglutide improved ALT (15/19), AST (10/17), GGT (8/13), fibrosis-4 index (F4) (7/8), hepatic steatosis index (SFI) (5/5), AST to platelet ratio (3/4), controlled attenuation parameter (CAP) (4/4), fatty liver index (3/3) and triglyceride-glucose index (T2D) (3/3) vs. baseline, but had mixed effects on liver stiffness measurement (2/5), total bilirubin (1/3) and type IV collagen (1/3).
- Tirzepatide improved AST vs. baseline and ALT, AST and GGT vs. GLP-1 RA (dulaglutide) (each 1/1).
- No studies comparing semaglutide and tirzepatide were identified.

Lipid Profile

- GLP-1 RA-based therapies had mixed effects on high-density lipoprotein cholesterol (HDL-C) (2/9), low-density lipoprotein cholesterol (LDL-C) (6/10), total cholesterol (TC) (4/10) and triglycerides (TG) (4/9) vs. baseline.
- Semaglutide improved LDL-C (35/45), non-HDL-C (4/5), TC (30/34) and TG (32/47) vs. baseline. Most studies reported no effect on HDL-C (8/45). Improvements were reported in LDL-C (2/2), with mixed results for HDL-C (1/2) and TG (1/2), vs. no GLP-1 RA.
- Tirzepatide improved HDL-C (2/5), LDL-C (5/7), TC (6/6) and TG (5/7) vs. baseline and TG (1/1) vs. no GLP-1 RA.
- No studies comparing semaglutide and tirzepatide were identified.

Figure 1. Evidence Map

- Neurological and Psychological Outcomes**
 - 17 unique studies
 - 4 global, 8 US, 2 Japan, 2 Spain, 1 Thailand
 - 48 different endpoints
- Bone Health**
 - 4 unique studies
 - 1 US, 1 Italy
 - 14 different endpoints
- Hepatic Outcomes**
 - 36 unique studies
 - 1 global, 10 Japan, 6 US, 8 Italy, 5 Spain, 3 UAE
 - 2 China, 2 UK, 1 Croatia, 1 Romania
 - 34 different endpoints
- Lipid Profile**
 - 73 unique studies
 - 2 global, 15 Italy, 13 Japan, 11 US, 6 Saudi Arabia, 5 Spain, 4 China, 4 UAE, 3 UK, 2 Germany, 1 Bangladesh, 1 Colombia, 1 Croatia, 1 India, 1 Pakistan, 1 Romania, 1 Slovenia, 1 Turkey
 - 19 different endpoints
- Cardiovascular**
 - 68 unique studies
 - 9 global, 17 US, 15 Italy, 7 Japan, 4 Saudi Arabia, 3 China, 3 Spain, 3 UK, 2 UAE, 1 Colombia, 1 Germany, 1 Pakistan, 1 Switzerland
 - 83 different endpoints
- Renal**
 - 44 unique studies
 - 3 global, 11 Italy, 11 US, 10 Japan, 7 Spain, 4 Saudi Arabia, 2 Columbia, 2 UAE, 2 UK, 1 China, 1 Croatia, 1 North Macedonia, 1 Romania
 - 18 different endpoints
- All-cause Mortality**
 - 15 unique studies
 - 11 global, 3 US, 1 Germany
 - 1 endpoint
- Cancer Risk**
 - 2 unique studies
 - 1 global, 1 US
 - 1 endpoint

Abbreviation: UAE = United Arab Emirates

Bone Health

- GLP-1 RAs improved bone turnover markers (carboxyterminal telopeptide [1/1], bone isoenzyme of alkaline phosphatase [1/1], adiponectin [1/1] and myostatin [1/1]) vs. baseline, but not bone mineral density (0/1). GLP-1 RAs reduced the risk of psoriasis (2/2) and osteoarthritis (1/1) and improved hypercalcaemia, hypocalcaemia, calcium and vitamin D (each 1/1) vs. no GLP-1 RA.
- Semaglutide reduced the risk of osteoarthritis (1/1) and hypercalcaemia (1/1) vs. no GLP-1 RA.
- Tirzepatide improved hypocalcaemia vs. no GLP-1 RA-based therapy (1/1).
- Tirzepatide reduced the risk of osteoarthritis vs. semaglutide (1/1).

Neurological and Psychological Outcomes

- GLP-1 RA-based therapies reduced alcohol consumption (1/1) and the risk of Alzheimer's disease (2/2), Lewy body dementia, vascular dementia and seizures (each 1/1) vs. no GLP-1 RA-based therapy. No effect on Parkinson's disease (0/1) was reported, and results for depression were mixed (1/2).
- Semaglutide improved cravings (2/2), emotional eating behaviours (2/2), binge episodes (1/1) and food addiction (1/1) vs. baseline. Semaglutide improved stimulant misuse, nicotine misuse, alcohol misuse, cannabis misuse and smoking cessation (each 3/3) vs. no GLP-1 RA. Semaglutide reduced the risk of depression, psychosis and suicidality vs. no GLP-1 RA, but had no impact on bipolar disorder, anxiety disorder, OCD or negative control outcomes (each 1/1). Semaglutide reduced the risk of Alzheimer's disease (2/2), encephalitis (1/1), cognitive deficit (1/1), dementia (1/1), epilepsy/seizures (1/1), insomnia (1/1), Lewy body dementia (1/1) and vascular dementia (1/1) vs. no GLP-1 RA.
- No studies were identified for tirzepatide.

Cancer Risk

- GLP-1 RAs (semaglutide, liraglutide, dulaglutide) reduced the risk of pancreatic, gastrointestinal, skin, breast, female genital, male genital, prostate, urinary tract, eye, brain, central nervous system, thyroid, respiratory, mesothelial and lymphoid/hematopoietic cancers (each 1/1) vs. no GLP-1 RA.
- No studies were identified for tirzepatide.

All-cause Mortality

- Both semaglutide (4/5) and tirzepatide (3/3) improved ACM vs. no GLP-1 RA-based therapy.
- Tirzepatide was associated with a lower risk of death vs. GLP-1 RAs (2/2) and semaglutide specifically (1/1).

Conclusions

RWE suggests that GLP-1 RA-based therapies have a broad range of benefits beyond weight loss and glycaemic control, which could be relevant to HTA bodies and healthcare systems when evaluating the health and economic value of these treatments. Of note, evidence suggests that GLP-1 RAs may have a significant positive impact on ACM when compared with no GLP-1 RA, and that semaglutide and tirzepatide improve a range of CV outcomes (including blood pressure), lipid profile and neurological/psychological outcomes compared with baseline. Further research is needed to understand the comparative effects of semaglutide and tirzepatide across all outcomes included in this review.

References

- European Medicines Agency. Wegovy (semaglutide). Summary of Product Characteristics. 2021. [European Medicines Agency](#). Accessed: 2025. [Wegovy \(semaglutide\). Summary of Product Characteristics. 2021. European Medicines Agency](#)
- European Medicines Agency. Mounjaro (tirzepatide). Summary of Product Characteristics. 2023. [European Medicines Agency](#). Accessed: 2025. [Mounjaro \(tirzepatide\). Summary of Product Characteristics. 2023. European Medicines Agency](#)
- Chen H, et al. *Diabetes Care*. 2024;47(12):2385-94. [Chen H, et al. Diabetes Care. 2024;47\(12\):2385-94.](#)
- Baranovskiy A, et al. *Diabetes Care*. 2024;47(12):2395-41. [Baranovskiy A, et al. Diabetes Care. 2024;47\(12\):2395-41.](#)
- Shah D, et al. *Diabetes Care*. 2024;47(12):2385-94. [Shah D, et al. Diabetes Care. 2024;47\(12\):2385-94.](#)
- Yano E, et al. *Diabetes Care*. 2025;48(1):127-37. [Yano E, et al. Diabetes Care. 2025;48\(1\):127-37.](#)
- Pearson-Stuttard J, et al. *Int J Obes*. 2023;47(12):1239-46. [Pearson-Stuttard J, et al. Int J Obes. 2023;47\(12\):1239-46.](#)

Disclosures

All authors are employees of PPD™ Evidera™ Health Economics and Market Access, Thermo Fisher Scientific. This poster was funded by Thermo Fisher Scientific. Editorial and graphic design support were provided by Anissa Tabouat and Heather Nashmore of Thermo Fisher Scientific.

Please scan this QR code with your smartphone camera or app to obtain the supplementary information. This information is available through digital download and for personal use only and may not be reproduced without permission from GSK/Novo. © is the authors of this poster.

98 | Conference Poster Portfolio

Conference Poster Portfolio | 99

Poster Title

Millennial Medical Record Data Profile: A Japanese Electronic Medical Records Database Utilising Unstructured Data for Lung Cancer Research

Objective

- The goal of this study was to describe the characteristics of MMR data, illustrating its utility for lung cancer research.

AUTHORS Ayumi Hamaguchi ¹ Kenichiro Shiba ² Mari Kato ³ Daisuke Wakasugi ⁴ Tadashi Koga ³ Suguru Nozue ⁴	AFFILIATION ¹ Thermo Fisher Scientific, London, UK ² General Incorporated Association Life Data Initiative, Kyoto, Japan ³ Clinical Study Support, Inc., Nagoya, Japan ⁴ NTT DATA Japan Corporation, Tokyo, Japan	PRACTICE AREA PPD™ Evidera™ Real-World Data & Scientific Solutions, Thermo Fisher Scientific	CLIENT INTERNAL General Incorporated Association Life Data Initiative Clinical Study Support, Inc NTT DATA Japan Corporation
---	--	--	---

RWD125

Millennial Medical Record Data Profile: A Japanese Electronic Medical Records Database Utilising Unstructured Data for Lung Cancer Research

Ayumi Hamaguchi,¹ Kenichiro Shiba,² Mari Kato,³ Daisuke Wakasugi,⁴ Tadashi Koga,³ Suguru Nozue,⁴ Takayuki Sawada,³ Tomoyo Morita,⁴ Amanda Pulfer,¹ Dimitra Lambrelli¹

¹Thermo Fisher Scientific, London, UK; ²General Incorporated Association Life Data Initiative, Kyoto, Japan; ³Clinical Study Support, Inc., Nagoya, Japan; ⁴NTT DATA Japan Corporation, Tokyo, Japan

Background

- Real-world data sources, such as health insurance claims and Diagnosis Procedure Combination (DPC) databases, are commonly used in Japanese research because of their broad coverage and suitability for longitudinal analyses,^{1,2} but these sources often lack detailed clinical information.
- Electronic medical records (EMRs) offer rich clinical detail, including laboratory values and indicators of disease severity and progression,³ but access to EMR-based data—particularly for commercial research—remains limited.
- The Millennial Medical Record (MMR) database includes data from EMRs, claims, and DPC sources. It has the potential to be leveraged for real-world research, including lung cancer and other disease areas, using its comprehensive clinical and longitudinal data.
- To describe the characteristics of the MMR database, lung cancer was selected due to its clinical relevance and its status in Japan as a leading cause of death,⁴ with its rising age-adjusted incidence rate since 2019.⁵

Objectives

- The goal of this study was to describe the characteristics of MMR data, illustrating its utility for lung cancer research.

Methods

MMR Overview

- The MMR database includes structured and unstructured data from EMRs (Table 1) that complies with the Medical Markup Language 4.2.0 standard,⁶ claims data, DPC data, and imaging data.
- Data are collected under the framework of the Next Generation Medical Infrastructure Act,⁷ which allows only certified entities to manage and anonymise medical information for research use. Processing staff complete mandatory training in anonymisation, data security, and compliance (Figure 1). The MMR database is operated by the Life Data Initiative and NTT DATA Japan Corporation.
- As MMR's coverage expands throughout Japan, the number of participating hospitals increased from 43 in 2020 to 67 in 2025 (Figure 2). Of these, the number contributing EMR, DPC, and claims data usable for research increased from 10 to 26 over the same period.
- Figure 3 provides an overview of data availability in the MMR database.

Category	Details
Participating hospitals	67 secondary/tertiary care hospitals with an average of 600 beds, mostly university and regional core hospitals under secondary use agreements (in LDJ as of Aug 2025)
Patient coverage	1.9 million patients (approx. 1.5% of Japanese population) have complete EMR + DPC + claims data, as of Feb 2025
Data collection	Started in 2016 and is ongoing*
Update frequency	Quarterly updates*
Data access	Review required by the LDJ (Purpose of Data Use Review Board)

Abbreviations: DPC = Diagnosis Procedure Combination; EMR = electronic medical record; LDJ = Life Data Initiative; MMR = Millennial Medical Record. *Data updated via automatic transfer or secure file upload from participating hospitals.

Figure 1. Data Processing and Anonymisation

- Out-of-procedure managed by contributing hospitals
- Removal of direct identifiers
- Creation of pseudonymised patient IDs for cross-source linkage between EMR, DPC, and claims
- Re-identification risk assessed; low-frequency or potentially identifying values suppressed if needed

Figure 2. Hospitals Contributing to MMR, by Region

Region	Total
Hokkaido	4
Tohoku	6
Kanto	12
Kansai	6
Tokai	3
Kinki (Kansai)	17
Chugoku	2
Shikoku	4
Nagasaki	20
Okinawa	0

Abbreviations: DPC = Diagnosis Procedure Combination; EMR = electronic medical record; LDJ = Life Data Initiative; MMR = Millennial Medical Record. *Data anonymisation must be performed in a secure workspace (e.g., fingerprint access, camera monitoring, no personal devices).

Figure 3. MMR Data Overview

Structured data	EMR data
<ul style="list-style-type: none">Demographics (sex, date of birth, nationality, insurance information, etc.)Diagnosis history (clinical codes/medications using ICD-10 codes, disease onset and end date, outcome, etc.)Prescriptions and injections (drug code, daily dosage, route, start date, end date, etc.)Laboratory tests (specimen type, IJAC10 code, test result, reference range, flag for abnormal value, etc.)Vital signs (temperature, blood pressure, pulse, oxygen saturation)	<ul style="list-style-type: none">Clinical notes (progress notes, physician communication)Discharge summaries (admission course, findings, outcomes)Radiology reportsReferral letters and consultation notes

DPC data

- Admission and discharge information
- DPC group information

Claims data

- Demographics information
- Prescriptions information
- Hospital costs
- Diagnoses
- Medical practice

Imaging data

- X-ray, CT, MRI

Abbreviations: DPC = Diagnosis Procedure Combination; MMR = electronic medical record; IJAC10 = Japan Laboratory Code, version 10; MMR = Millennial Medical Record.

Statistical Analysis

- To describe MMR data characteristics, age distribution as of 2024 was calculated and compared with the National Database of Health Insurance Claims and Specific Health Checkups of Japan (NDB),⁸ using patients with a first-visit claim between April 2023 and March 2024.
- Patients with lung cancer were identified using ICD-10 code a C34* as of February 2025, with counts stratified by sex and age.
- To demonstrate the value of EMR data, keywords were grouped into five domains: histology, genetic tests, disease scores, metastasis sites, and tumour marker tests. A rule-based keyword search was applied to assess data availability, regardless of confirmed presence. Keywords for histology, genetic tests, and disease scores were extracted from unstructured data using SQL scripts program supported by Generative AI. Analyses were descriptive.

Results

- As of 2024, the MMR database included 1.6 million individuals from 26 hospitals contributing EMR, DPC, and claims data across Japan, while the NDB database included 88 million individuals. Overall age distributions were similar, with all proportional differences under 5%, supporting the MMR database's representativeness. The largest gap was 4.45% in the 75–79 age group (Figure 4).

Figure 4. Age Distribution Among MMR and NDB

Abbreviations: MMR = Millennial Medical Record; NDB = National Database of Health Insurance Claims and Specific Health Checkups of Japan. Note: The NDB was used as a reference due to its comprehensive nationwide coverage.

- Among 29,545 patients with lung cancer, 63.1% were male (Figure 5). The largest age group was 70–79 years old (44.2%), consistent with national incidence data showing higher occurrence in adults aged 60 and older than in younger age groups.⁹

Figure 5. Age and Sex Distribution of Patients With Lung Cancer in MMR

Abbreviation: MMR = Millennial Medical Record.

- Among 29,545 patients with lung cancer, keywords indicating tissue types—adenocarcinoma, small cell, and squamous cell carcinoma—were identified in 10%–40% of cases from unstructured data (Table 2). The keyword list was indicative; clinician validation would expand data identification.

Table 2. Availability of Clinical Data in Unstructured and Structured EMR

Variables	Keywords Lists/Description	Percentage (%) of Patients With Keywords/Codes
Histology		
Adenocarcinoma	(Adenocarcinoma or AdE)-(肺)-(野)-without negative expressions*, OR (肺)(腺癌 or 腺状腺癌)-without negative expressions*	35–40%
Small cell carcinoma	(Small Cell Lung Cancer or SCLC)-without negative expressions*, OR (肺)(小細胞癌)-without negative expressions*, OR (肺)(小細胞)-肺)-without negative expressions*	10–15%
Squamous cell carcinoma	(Squamous Cell Carcinoma or SCC)-(肺 or 肺); OR (肺)(扁平上皮)-(肺)-without negative expressions*	10–15%
Genetic test		
EGFR	EGFR†	20–25%
ALK	ALK†	15–20%
ROS1	ROS1†	5–10%
Disease score		
Performance status	(PS)-Any of the numbers 0–2; OR (Performance)-(Status)-any of the numbers 0–9	55–60%
Metastasis site		
Brain	(骨 or Brain)-(転移 or 病変 or Meta)-without negative expressions*	Keywords shown for reference purposes
Lymph node	(リンパ)-転移 or 病変 or Meta)-without negative expressions*	
Data from structured EMR		
Tumour marker tests		
CEA	–	70–75%
SCC	–	35–40%
CYFR21-1	–	15–20%

Abbreviations: ALK = anaplastic lymphoma kinase; CEA = carcinoembryonic antigen; CYFR21-1 = cytokeratin 19 fragment; EGFR = epidermal growth factor receptor; PS = performance status; SCC = squamous cell carcinoma; SCLC = squamous cell carcinoma lung cancer; SCLC = small cell lung cancer. *Negative expressions indicating absence or uncertainty of a specific condition in Japanese include (無), (欠), (無), (無), (無), or (無). †Any alphabet before EGFR/ALK/ROS1 are excluded.

Conclusions

- Age distributions were generally similar between the MMR and NDB databases, supporting its representativeness.
- By integrating complete EMR data—including histological type and biomarkers from unstructured records—with claims and DPC data, the MMR database enables research closely aligned with routine clinical practice, creating opportunities for real-world studies that complement clinical trials.
- The growth of the MMR database demonstrates the feasibility of harnessing EMR data for large-scale research in Japan. As coverage expands, the MMR will support innovative real-world evidence generation—including external control arms and longitudinal studies—ultimately accelerating advances in treatment and health policy.

References: 1. Asia Clin Oncol. 2023;9(10):644. 2. Fujimori J, Fujimori J, Doori T. Disparities in Health Outcomes. 2023;9(10):650–3. 3. Zhou Y, Tsubota T. MMR. Med. 2023;11:e0080. 4. World Health Organization. Global Cancer Observatory. Japan. 2020. 5. Japan Cancer Statistics Working Group. Cancer Statistics in Japan 2023. 6. Medical Markup Language 4.2.0. 7. Next Generation Medical Infrastructure Act. 8. Tomiyama H, et al. J Clin Oncol. 2023;41(16):2468–75. 9. Japanese Cancer Statistics Working Group. Cancer Statistics in Japan 2023. 10. Japanese Cancer Statistics Working Group. Cancer Statistics in Japan 2023. 11. Japanese Cancer Statistics Working Group. Cancer Statistics in Japan 2023. 12. Japanese Cancer Statistics Working Group. Cancer Statistics in Japan 2023. 13. Japanese Cancer Statistics Working Group. Cancer Statistics in Japan 2023. 14. Japanese Cancer Statistics Working Group. Cancer Statistics in Japan 2023. 15. Japanese Cancer Statistics Working Group. Cancer Statistics in Japan 2023. 16. Japanese Cancer Statistics Working Group. Cancer Statistics in Japan 2023. 17. Japanese Cancer Statistics Working Group. Cancer Statistics in Japan 2023. 18. Japanese Cancer Statistics Working Group. Cancer Statistics in Japan 2023. 19. Japanese Cancer Statistics Working Group. Cancer Statistics in Japan 2023. 20. Japanese Cancer Statistics Working Group. Cancer Statistics in Japan 2023. 21. Japanese Cancer Statistics Working Group. Cancer Statistics in Japan 2023. 22. Japanese Cancer Statistics Working Group. Cancer Statistics in Japan 2023. 23. Japanese Cancer Statistics Working Group. Cancer Statistics in Japan 2023. 24. Japanese Cancer Statistics Working Group. Cancer Statistics in Japan 2023. 25. Japanese Cancer Statistics Working Group. Cancer Statistics in Japan 2023. 26. Japanese Cancer Statistics Working Group. Cancer Statistics in Japan 2023. 27. Japanese Cancer Statistics Working Group. Cancer Statistics in Japan 2023. 28. Japanese Cancer Statistics Working Group. Cancer Statistics in Japan 2023. 29. Japanese Cancer Statistics Working Group. Cancer Statistics in Japan 2023. 30. Japanese Cancer Statistics Working Group. Cancer Statistics in Japan 2023. 31. Japanese Cancer Statistics Working Group. Cancer Statistics in Japan 2023. 32. Japanese Cancer Statistics Working Group. Cancer Statistics in Japan 2023. 33. Japanese Cancer Statistics Working Group. Cancer Statistics in Japan 2023. 34. Japanese Cancer Statistics Working Group. Cancer Statistics in Japan 2023. 35. Japanese Cancer Statistics Working Group. Cancer Statistics in Japan 2023. 36. Japanese Cancer Statistics Working Group. Cancer Statistics in Japan 2023. 37. Japanese Cancer Statistics Working Group. Cancer Statistics in Japan 2023. 38. Japanese Cancer Statistics Working Group. Cancer Statistics in Japan 2023. 39. Japanese Cancer Statistics Working Group. Cancer Statistics in Japan 2023. 40. Japanese Cancer Statistics Working Group. Cancer Statistics in Japan 2023. 41. Japanese Cancer Statistics Working Group. Cancer Statistics in Japan 2023. 42. Japanese Cancer Statistics Working Group. Cancer Statistics in Japan 2023. 43. Japanese Cancer Statistics Working Group. Cancer Statistics in Japan 2023. 44. Japanese Cancer Statistics Working Group. Cancer Statistics in Japan 2023. 45. Japanese Cancer Statistics Working Group. Cancer Statistics in Japan 2023. 46. Japanese Cancer Statistics Working Group. Cancer Statistics in Japan 2023. 47. Japanese Cancer Statistics Working Group. Cancer Statistics in Japan 2023. 48. Japanese Cancer Statistics Working Group. Cancer Statistics in Japan 2023. 49. Japanese Cancer Statistics Working Group. Cancer Statistics in Japan 2023. 50. Japanese Cancer Statistics Working Group. Cancer Statistics in Japan 2023. 51. Japanese Cancer Statistics Working Group. Cancer Statistics in Japan 2023. 52. Japanese Cancer Statistics Working Group. Cancer Statistics in Japan 2023. 53. Japanese Cancer Statistics Working Group. Cancer Statistics in Japan 2023. 54. Japanese Cancer Statistics Working Group. Cancer Statistics in Japan 2023. 55. Japanese Cancer Statistics Working Group. Cancer Statistics in Japan 2023. 56. Japanese Cancer Statistics Working Group. Cancer Statistics in Japan 2023. 57. Japanese Cancer Statistics Working Group. Cancer Statistics in Japan 2023. 58. Japanese Cancer Statistics Working Group. Cancer Statistics in Japan 2023. 59. Japanese Cancer Statistics Working Group. Cancer Statistics in Japan 2023. 60. Japanese Cancer Statistics Working Group. Cancer Statistics in Japan 2023. 61. Japanese Cancer Statistics Working Group. Cancer Statistics in Japan 2023. 62. Japanese Cancer Statistics Working Group. Cancer Statistics in Japan 2023. 63. Japanese Cancer Statistics Working Group. Cancer Statistics in Japan 2023. 64. Japanese Cancer Statistics Working Group. Cancer Statistics in Japan 2023. 65. Japanese Cancer Statistics Working Group. Cancer Statistics in Japan 2023. 66. Japanese Cancer Statistics Working Group. Cancer Statistics in Japan 2023. 67. Japanese Cancer Statistics Working Group. Cancer Statistics in Japan 2023. 68. Japanese Cancer Statistics Working Group. Cancer Statistics in Japan 2023. 69. Japanese Cancer Statistics Working Group. Cancer Statistics in Japan 2023. 70. Japanese Cancer Statistics Working Group. Cancer Statistics in Japan 2023. 71. Japanese Cancer Statistics Working Group. Cancer Statistics in Japan 2023. 72. Japanese Cancer Statistics Working Group. Cancer Statistics in Japan 2023. 73. Japanese Cancer Statistics Working Group. Cancer Statistics in Japan 2023. 74. Japanese Cancer Statistics Working Group. Cancer Statistics in Japan 2023. 75. Japanese Cancer Statistics Working Group. Cancer Statistics in Japan 2023. 76. Japanese Cancer Statistics Working Group. Cancer Statistics in Japan 2023. 77. Japanese Cancer Statistics Working Group. Cancer Statistics in Japan 2023. 78. Japanese Cancer Statistics Working Group. Cancer Statistics in Japan 2023. 79. Japanese Cancer Statistics Working Group. Cancer Statistics in Japan 2023. 80. Japanese Cancer Statistics Working Group. Cancer Statistics in Japan 2023. 81. Japanese Cancer Statistics Working Group. Cancer Statistics in Japan 2023. 82. Japanese Cancer Statistics Working Group. Cancer Statistics in Japan 2023. 83. Japanese Cancer Statistics Working Group. Cancer Statistics in Japan 2023. 84. Japanese Cancer Statistics Working Group. Cancer Statistics in Japan 2023. 85. Japanese Cancer Statistics Working Group. Cancer Statistics in Japan 2023. 86. Japanese Cancer Statistics Working Group. Cancer Statistics in Japan 2023. 87. Japanese Cancer Statistics Working Group. Cancer Statistics in Japan 2023. 88. Japanese Cancer Statistics Working Group. Cancer Statistics in Japan 2023. 89. Japanese Cancer Statistics Working Group. Cancer Statistics in Japan 2023. 90. Japanese Cancer Statistics Working Group. Cancer Statistics in Japan 2023. 91. Japanese Cancer Statistics Working Group. Cancer Statistics in Japan 2023. 92. Japanese Cancer Statistics Working Group. Cancer Statistics in Japan 2023. 93. Japanese Cancer Statistics Working Group. Cancer Statistics in Japan 2023. 94. Japanese Cancer Statistics Working Group. Cancer Statistics in Japan 2023. 95. Japanese Cancer Statistics Working Group. Cancer Statistics in Japan 2023. 96. Japanese Cancer Statistics Working Group. Cancer Statistics in Japan 2023. 97. Japanese Cancer Statistics Working Group. Cancer Statistics in Japan 2023. 98. Japanese Cancer Statistics Working Group. Cancer Statistics in Japan 2023. 99. Japanese Cancer Statistics Working Group. Cancer Statistics in Japan 2023. 100. Japanese Cancer Statistics Working Group. Cancer Statistics in Japan 2023.

Presented at ISPOR —The Professional Society for Health Economics and Outcomes Research—Europe | 9–12 November 2025 | Glasgow, Scotland, UK

100 | Conference Poster Portfolio

Conference Poster Portfolio | 101


Poster Title

AI-Assisted Time-to-Event Projection: A Case Study and Broader Potential

Objective

- We explored the use of large language models (LLMs) (e.g., Chat-GPT, Gemini) to assist with understanding the prognosis of the study population to establish plausible ranges of life expectancy or median survival times.

AUTHORS K. Jack Ishak J. Jaime Caro	AFFILIATION Thermo Fisher Scientific, Montreal, Canada	PRACTICE AREA PPD™ Evidera™ Health Economics & Market Access, Thermo Fisher Scientific	CLIENT INTERNAL
--	--	--	---------------------------



AI-Assisted Time-to-Event Projection: A Case Study and Broader Potential

K. Jack Ishak,¹ J. Jaime Caro¹
¹Thermo Fisher Scientific, Montreal, Canada

Background

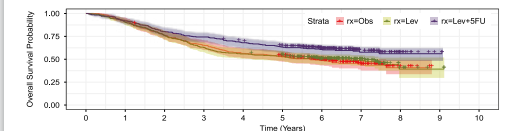
- Time-to-event projection involves balancing clinical information about the disease or treatments being studied and statistical fit considerations to ensure the selected model is appropriate.
- We explored the use of large language models (LLMs) (e.g., Chat-GPT, Gemini) to assist with understanding the prognosis of the study population to establish plausible ranges of life expectancy or median survival times.

Case Study: Overall Survival in Stage B/C Colon Cancer

Data Source and Analysis

- Randomized trial demonstrating benefit of adjuvant chemotherapy for colon cancer!
- Compared levamisole (Lev) and levamisole with 5-FU (Lev+5FU) to observation alone (Obs)
- Outcome of interest: overall survival (OS) (Figure 1)

Figure 1. Observed time-to-death distribution with Lev+5FU, Lev, and Obs

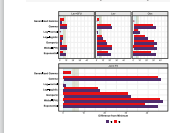
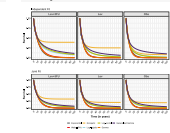
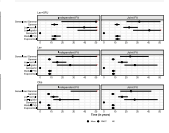


Abbreviations: Lev = levamisole; Lev+5FU = levamisole with 5FU; Obs = observation alone

Parametric Fitting Analyses

- Standard parametric analyses with supporting diagnostic assessments were carried out to select an optimal fit.
- Analyses considered joint fitting (i.e., assuming proportional hazards) and separate fitting with exponential, Weibull, Gompertz, log-logistic, log-normal, gamma, and generalized gamma distributions.
- Results from analyses suggest hazards are not proportional over time, and log-normal and generalized gamma provide best statistical fit and produce projections that are middle of the road (i.e., between more aggressive estimates from Weibull, for example, and implausibly optimistic fits from Gompertz).
- Selected results from fitting analyses are shown in Figure 2.

Figure 2. Fit statistics, projections and predicted mean event times from projection analyses

Fit statistics (AIC and BIC)	Projected time-to-death curves	Predicted mean times to death
		

*Point estimate, lower bound or upper bound not estimated.
RMST calculated up to 90th percentile of the fitted distribution.
Abbreviations: AIC = Akaike Information Criterion; BIC = Bayesian Information Criterion; Lev = levamisole; Lev+5FU = levamisole with 5FU; NE = not estimated; Obs = observation alone; PH = proportional hazard; RMST = restricted mean survival time.

AI as a Clinical Assistant

- We iteratively prompted the LLM to provide and request information as described below (Figure 3).
- The LLM provided detailed responses at each step, including relevant cautionary/advisory points, such as:
 - The current applicability of the observed results (since the data are from a study conducted pre-1990s)
 - Looking for convergence to general mortality after 10–15 years “as eventually, the excess cancer mortality risk should diminish.”
 - Possibility of functional cure supporting 25+ years of maximum survival
- The LLM selected the jointly fitted log-normal distribution based on fit, plausibility, and stability of predictions (e.g., mean survival) and suggested sensitivity analyses to deal with limitations of selected fits (i.e., optimistic tails).

Figure 3. Overview of information exchange with LLM to obtain insights to guide model selection

Context & Background Information	A Priori Expected Survival	Interpretation and Recommendations	Dealing with Implausible Tails	Summary
<ul style="list-style-type: none">Assign persona (clinical expert advising analysis)Ask to review the study being analyzedAsk to summarize the study population and clinical profile	<ul style="list-style-type: none">Ask for estimates of the expected survival of participants of the study, considering the period in which the study is conductedProvide characteristics of population (i.e. male, age, etc.) which fit to select for use in an economic model	<ul style="list-style-type: none">Provide a structured summary, including predicted mean, median and maximum survival times, best statistical fit, and those with comparable or similar fit (AIC/BIC within 4.7 points)Ask for recommendations on which fit to select for use in an economic model	<ul style="list-style-type: none">Ask for suggestions for how to deal with implausible predictions or tails from the selected distribution	<ul style="list-style-type: none">Ask for a summary of the background recommendations and suggestions for handling limitations of fits

Abbreviations: AIC = Akaike Information Criterion; BIC = Bayesian Information Criterion; LLM = large language models

Insights about Population Profile

- The LLM characterized the population as follows:
 - In short, the cohort you are modeling is a high-risk but relatively fit population that had undergone complete removal of their primary tumor, with nodal involvement being the dominant driver of their poor long-term prognosis.
 - When extrapolating the survival tail of your model, be mindful that non-cancer mortality related to toxicity/poor compliance might have been higher in the older cohort than it would be today.
- It cautioned about bias and applicability of results from the study, and considerations for projection analyses.

A Priori Expected Survival

- The LLM estimated the median survival to be around 3.5–4.5 years and noted that long-term survivors would pull the mean (expected) survival to 5–7 years.
- It suggested survival could extend to 25 years or more in some patients since “a small but significant fraction of patients are cured by surgery alone” and a potentially young age at diagnosis in some patients.

Interpretation and Recommendations

- It provided a recommended fit with rationale explaining the selection (as well as rationale for excluded models – not shown) and suggested sensitivity analyses with alternative distributions.

Metric	Log-normal (Joint Fit)	Rationale
Goodness-of-Fit	Best fit according to the BIC and competitive with AIC; best fit (Generalized Gamma). BIC penalizes models with more parameters, favoring parsimony – a desirable trait for extrapolation.	
Predicted Median OS	4.18 years (I: 4.0–9.93)	This is clinically plausible for an effective adjuvant regimen in this high-risk population.
Predicted Mean OS	22.62 years (I: 7.42–30.39)	While high, this is a finite, numerically stable estimate and is far more reasonable than the highly inflated means from the Generalized Gamma and Log-logistic models.
Long-Term Plausibility	10% event-free at 50.87 years.	This extrapolation, though optimistic, is contained within a clinically plausible timeframe for a proportion of younger patients to survive.

Dealing with Implausible Fits

- The LLM’s assessment that the tail of jointly fitted log-normal is “optimistic” was questionable, particularly the 90th percentile stretched to hundreds of years.
- When prompted for recommendations on how to best deal with the implausible tails, the LLM proposed three strategies with supporting explanations to implement these:
 - Blending with General Population Mortality
 - Imposing a “Cure Point” (Curtailment)
 - Restricting the Mean Survival Time (RMST)

Summary of Insights

- The LLM produced a report of the insights requested during the exchange and a summary of key points.

Conclusions

- This case study shows that LLMs can be helpful in providing clinical insights to aid analysts to interpret results from time-to-event fitting.
- The LLM demonstrated an understanding of considerations for projecting time-to-event curves (e.g., behavior of long tails and impact on mean vs. median OS) and flagged relevant clinical issues on use of the data.
- Responses from the LLM on a priori expected survival were particularly useful as this type of information is not always readily available to analysts.
 - The LLM’s responses were supported with explanations of its reasoning (e.g., risk factors driving estimates).
 - The LLM produced a clear summary of its recommendations, capturing the salient points from the exchange.
- Responses from the LLM may be influenced by the extent of information provided and how this is structured.
 - We summarized results to provide to the LLM for interpretation in which we identified best-fitting models (based on AIC/BIC) and labeled others as comparable or similar, as appropriate.
 - Changing this format and the type of information provided could affect responses; further exploration of this would be beneficial.
- While this case study focused on clinical insights and interpretation, LLMs can also help with suggestions of analyses to perform, coding support, and communication of results.

References

1. CG Moerliet, et al. *N Engl J Med*. 332:352-358, 1990.

Disclosures


All authors are employees of PPD™ Evidera™ Health Economics & Market Access, Thermo Fisher Scientific. The development of this poster was funded by Thermo Fisher Scientific.

Acknowledgments

Editorial and graphic design support were provided by Sean Smith and Shari Berger of Thermo Fisher Scientific.

Presented at ISPOR – The Professional Society for Health Economics and Outcomes Research—Europe | 9–12 November 2025 | Glasgow, Scotland, UK

© 2025 Thermo Fisher Scientific Inc. All rights reserved. All trademarks are the property of Thermo Fisher Scientific and its subsidiaries unless otherwise specified. This information is not intended to encourage use of these products by any receiver that might infringe the intellectual property rights of others.





ISPOR EU 2025

9–12 November 2025 | Glasgow, Scotland, UK


Poster Title

Have Delays in Vaccine Market Access Improved in Europe and the US Over the Last Decade?

Objective

- To determine if delays between vaccine licensure and National Immunization Technical Advisory Group (NITAG) evaluation have improved over the last decade.

AUTHORS Samantha James ¹ Anna D'Ausilio ² Jennifer Boss ³ Ruth Chapman ³ Kassandra Schaible ⁴ Aleksandra Zapala Szufel ⁵	AFFILIATION ¹ Thermo Fisher Scientific, Paris, France ² Thermo Fisher Scientific, Milan, Italy ³ Thermo Fisher Scientific, London, UK ⁴ Thermo Fisher Scientific, Waltham, MA, USA ⁵ Thermo Fisher Scientific, Alicante, Spain	PRACTICE AREA PPD™ Evidera™ Health Economics & Market Access, Thermo Fisher Scientific	CLIENT INTERNAL
---	---	--	---------------------------


HTA170

Have Delays in Vaccine Market Access Improved in Europe and the US Over the Last Decade?

Samantha James,¹ Anna D'Ausilio,² Jennifer Boss,³ Ruth Chapman,³ Kassandra Schaible,⁴ Aleksandra Zapala Szufel⁵
¹Thermo Fisher Scientific, Paris, France; ²Thermo Fisher Scientific, Milan, Italy; ³Thermo Fisher Scientific, London, UK; ⁴Thermo Fisher Scientific, Waltham, MA, USA; ⁵Thermo Fisher Scientific, Alicante, Spain

Background

- The pathway to market access for vaccines differs from that for pharmaceuticals and varies significantly across countries and is dependent on the involvement and processes of National Immunization Technical Advisory Groups (NITAGs).
- The non-standard approach from one country to another creates delays between vaccine licensure and recommendations, which further delays consumer access to these imperative public health technologies.
- Previous analysis of vaccine market access indicated that, overall, the time from licensure to published decision was shorter for the US than in Europe. However, the delay was greater than that observed for non-vaccines.
- COVID-19 market access was rapid in comparison to previous vaccine assessments. To investigate if this has accelerated the market access for other vaccines, we examined the time from licensure to recommendation in France, Ireland, Italy, Poland, the UK, and the US.

Objective

- To determine if delays between vaccine licensure and National Immunization Technical Advisory Group (NITAG) evaluation have improved over the last decade.

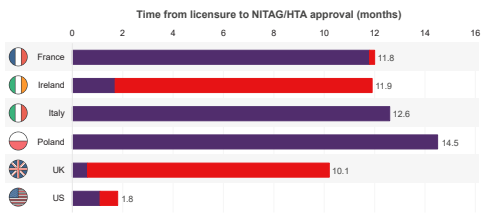
Methods

- A literature review was conducted to determine the time from licensure to NITAG/health technology assessment (HTA) recommendation for vaccines approved by the EMA, FDA and the Medicines and Healthcare products Regulatory Agency (MHRA) in the UK for the prevention of respiratory syncytial virus (RSV).
- Six countries (France, Ireland, Italy, Poland, the UK, and the US) were selected based on their decision-making, funding, and tendering archetypes.
- Delays for RSV vaccines were compared with published delays for pneumococcal, human papillomavirus, and quadrivalent influenza vaccines.

Results

- Three RSV vaccines have been approved for use in older adults (Abyryo [Pfizer Inc.], Arexvy [GlaxoSmithKline Biologicals] and mResvia [Moderna Inc.]), younger adults (Abyryo), high-risk younger adults (Arexvy and mResvia), and pregnant women (Abyryo).
- Vaccines for pregnant women were reviewed by NITAG/HTA in all countries except Poland (Figure 1). Vaccines for older adults (≥60 years of age) were reviewed by NITAG/HTA in all countries except mResvia in Ireland (Figure 2). No RSV vaccines were reviewed by NITAG/HTA for use in younger adults, except in Italy where an HTA decision was made within 3 months for the use of Abyryo in adults 18–59 years of age, and within 10.3 months for the use of Arexvy in high-risk adults 50–59 years of age.
- NITAG/HTA recommendations for RSV vaccines were made within a mean of 8.5–13.9 months of licensure in Europe (8.5, 10.2, 10.7, 12.8, and 13.9 in Italy, France, the UK, Ireland, and Poland, respectively) versus 2.2 months in the US (Figure 3).
- In comparison, NITAG recommendations for earlier vaccines were made within 10 months of licensure in the US,¹ 2–6 years in Poland and Ireland,² and more than 6 years in France, Italy, and the UK.³

Figure 1. Time from licensure to NITAG/HTA recommendation for RSV vaccines licensed for use among pregnant women^a by country



Country	Time from licensure to NITAG/HTA approval (months)
France	11.8
Ireland	11.9
Italy	12.6
Poland	14.5
UK	10.1
US	1.8

^a Additional time to CEESP opinion (France), HGA recommendation (Ireland), NITAG recommendation (Italy), ACTMAT recommendation (Poland), government decision (UK), and CDC publication (US). Time to CT opinion (France), NIAC recommendation (Ireland), AIFA decision (Italy), GIS recommendation (Poland), JCVI recommendation (UK) and ACIP recommendation (US) from EMA, FDA or MHRA approval.

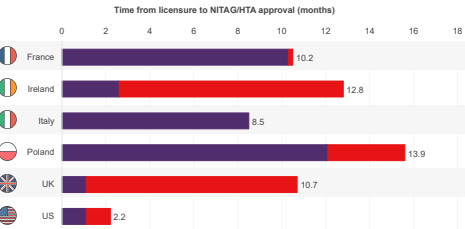
Figure 2. Time from licensure to NITAG/HTA recommendation for RSV vaccines licensed for use among older adults (≥60 years of age) by country



Country	Vaccine	Time from licensure to NITAG/HTA approval (months)
France	Abyryo (Pfizer Inc.)	12.4
	Arexvy (GlaxoSmithKline Biologicals)	15.0
	mResvia (Moderna Inc.)	2.1
Ireland	Abyryo (Pfizer Inc.)	11.9
	mResvia (Moderna Inc.)	Not yet evaluated by NITAG/HTA
Italy	Abyryo (Pfizer Inc.)	12.6
	Arexvy (GlaxoSmithKline Biologicals)	5.9
	mResvia (Moderna Inc.)	6.3
Poland	Abyryo (Pfizer Inc.)	14.9
	Arexvy (GlaxoSmithKline Biologicals)	23.8
	mResvia (Moderna Inc.)	2.4
UK	Abyryo (Pfizer Inc.)	10.1
	Arexvy (GlaxoSmithKline Biologicals)	11.4
US	Abyryo (Pfizer Inc.)	1.7
	Arexvy (GlaxoSmithKline Biologicals)	2.6
	mResvia (Moderna Inc.)	2.5

■ Additional time to CEESP opinion (France), HGA recommendation (Ireland), NITAG recommendation (Italy), ACTMAT recommendation (Poland), government decision (UK), and CDC publication (US). ■ Time to CT opinion (France), NIAC recommendation (Ireland), AIFA decision (Italy), GIS recommendation (Poland), JCVI recommendation (UK) and ACIP recommendation (US) from EMA, FDA or MHRA approval.

Figure 3. Mean time from licensure to NITAG/HTA recommendation for all licensed RSV vaccines by country



Country	Mean time from licensure to NITAG/HTA approval (months)
France	10.2
Ireland	12.8
Italy	8.5
Poland	13.9
UK	10.7
US	2.2

■ Additional time to CEESP opinion (France), HGA recommendation (Ireland), NITAG recommendation (Italy), ACTMAT recommendation (Poland), government decision (UK), and CDC publication (US). ■ Time to CT opinion (France), NIAC recommendation (Ireland), AIFA decision (Italy), GIS recommendation (Poland), JCVI recommendation (UK) and ACIP recommendation (US) from EMA, FDA or MHRA approval.

Conclusions

- The time between vaccine licensure and NITAG recommendation has dramatically improved in Europe and the US over the last decade.
- Nevertheless, a substantial discrepancy remains, with delays being four to nine times longer in Europe than in the US.

References

- James S, et al. *Value Health*. 2025;28(6):S264. 2. Laigle V, et al. *Vaccine*. Sep 15 2021;39(39):5706-5718.

Disclosures


SJ, AD, JB, RC, KS, and AZS are employees of PPD™ Evidera™ Health Economics & Market Access, Thermo Fisher Scientific. This poster was funded by Thermo Fisher Scientific.

Acknowledgments

Editorial and graphic design support were provided by Caroline Cole, Linda Allen, and Richard Leason of Thermo Fisher Scientific.

Presented at ISPOR—The Professional Society for Health Economics and Outcomes Research—Europe | 9–12 November 2025 | Glasgow, Scotland, UK

© 2025 Thermo Fisher Scientific Inc. All rights reserved. All trademarks are the property of Thermo Fisher Scientific and its subsidiaries unless otherwise specified. This information is not intended to encourage use of these products in any manner that might infringe the intellectual property rights of others.



Poster Title

NICE Has Embraced Single-arm Trials in Technology Appraisals, but Where is the Guidance?

Objective

- This study aimed to identify published guidance by National Institute for Health and Care Excellence (NICE) on SAT use in reimbursement submissions and to evaluate the acceptance of technology appraisals (TA) that use evidence solely from SATs across different therapeutic areas.

AUTHORS

Eleni Karinou
Charlotte Mitchell
Kacey Rawson

AFFILIATION

Thermo Fisher Scientific, London, UK

PRACTICE AREA

PPD™ Evidera™ Health Economics & Market Access, Thermo Fisher Scientific

CLIENT

INTERNAL

NICE Has Embraced Single-arm Trials in Technology Appraisals, but Where is the Guidance?

Eleni Karinou,¹ Charlotte Mitchell,¹ Kacey Rawson¹
¹Thermo Fisher Scientific, London, UK



HTA254

Background

Guidance and consistency in decision making by health technology assessment (HTA) agencies on the use of single-arm trials (SAT) as part of appraisals are being outpaced by the deployment of SAT methodology beyond advanced-stage cancer and rare disease. SATs are often used when randomised controlled trials (RCTs) are unfeasible, such as in small patient populations, rapidly evolving treatment landscapes, or when withholding treatment would be considered unethical.¹

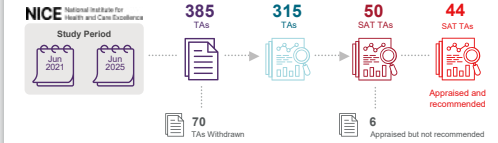
Objectives

This study aimed to identify published guidance by National Institute for Health and Care Excellence (NICE) on SAT use in reimbursement submissions and to evaluate the acceptance of technology appraisals (TA) that use evidence solely from SATs across different therapeutic areas.

Methods

We searched the NICE website for guidance on the use of SATs as sufficient evidence for decision-making. NICE TAs published between June 2021, and June 2025 were screened (Figure 1).²⁻¹¹ We counted the number of indications with a SAT-based submission and assessed their indirect comparison methodologies.

Figure 1. Methodology



Results

In each successive year, the proportion of all NICE TAs based on SATs increased (12% in the second half of June 2021 to 20% in 2024). While oncology remained the predominant area of all SAT-based submissions between June 2021 and June 2025 (72%), a number of SAT-based TAs in other disease areas were identified, including genetic disorders (14%), blood disorders (10%), and kidney disease (4%) (Figure 2).

Figure 2. The proportion of NICE TAs based on SATs has gradually increased over the past 5 years

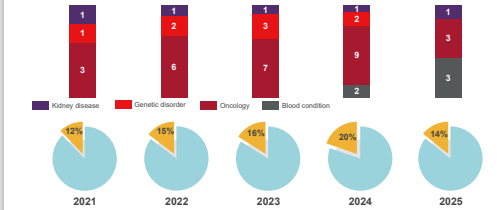
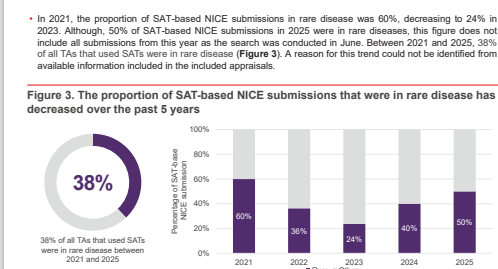


Figure 3. The proportion of SAT-based NICE submissions that were in rare disease has decreased over the past 5 years

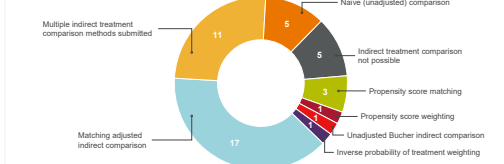


Abbreviations: NICE = National Institute for Health and Care Excellence; SAT = single arm trial; TA = technology appraisal. Note: Data for 2021 and 2022 are incomplete due to June search limits. Based on appraisals for 2023 to date, more SAT-based appraisals are expected from June through to December.

Results (cont.)

A total of eight different statistical methods to generate external comparator data were used across the included appraisals. Matching-adjusted indirect comparison was the most commonly used method to generate external comparator data (30%; 17 of 44 appraisals). A substantial proportion of TAs also submitted multiple indirect treatment comparisons (25%; 11 of 44 appraisals) (Figure 4).

Figure 4. MAIC was the most commonly used approach for generating comparative efficacy in the SAT-based NICE appraisals which received a positive recommendation (N=44)



Abbreviations: NICE = National Institute for Health and Care Excellence; RCT = randomised controlled trial; SAT = single-arm trial. There is no clear guidance on the acceptable level of uncertainty in the comparative evidence base. In 83% of rejected appraisals, the NICE committee quoted uncertainty in the comparative evidence base as a reason for a negative recommendation. In 52% of recommended appraisals, the NICE committee still quoted uncertainty in the comparative evidence base in their decision-making (Figure 5).

Figure 5. A high proportion of SAT-based TAs discussed uncertainty in the comparative evidence base as a key factor for decision making in both recommended and rejected appraisals



Abbreviations: NICE = National Institute for Health and Care Excellence; SAT = single arm trial; TA = technology appraisal.

Conclusions

- Our research highlights the high acceptance rate of SAT-based submissions by NICE in a wide range of therapeutic areas beyond oncology and rare disease, indicating possible scope creep of SATs beyond rare diseases.
- In many of the TAs receiving a positive recommendation, uncertainty in the comparative evidence methods was raised but often unresolved in publicly available sources.
- Equally, in TAs receiving a negative recommendation, uncertainty in the comparative evidence methods was often listed as the primary reason. This data highlights a "double standard" in the acceptance of SATs by NICE.
- Consistency and transparency in the appraisal process are not possible without guidance on acceptable methods for incorporating SAT evidence.

References

1. European Medicines Agency. Single-arm trials as pivotal evidence for the authorization of medicines in the EU. <https://www.ema.europa.eu/en/medicines/clinical-trials/clinical-trials-evidence/clinical-trials-evidence-4>. 2. National Institute for Health and Care Excellence (NICE). Evidence-based medicine for health and care. <https://www.nice.org.uk/guidance/ta254>. 3. NICE. Evidence-based medicine for health and care. <https://www.nice.org.uk/guidance/ta254>. 4. NICE. Evidence-based medicine for health and care. <https://www.nice.org.uk/guidance/ta254>. 5. NICE. Evidence-based medicine for health and care. <https://www.nice.org.uk/guidance/ta254>. 6. NICE. Evidence-based medicine for health and care. <https://www.nice.org.uk/guidance/ta254>. 7. NICE. Evidence-based medicine for health and care. <https://www.nice.org.uk/guidance/ta254>. 8. NICE. Evidence-based medicine for health and care. <https://www.nice.org.uk/guidance/ta254>. 9. NICE. Evidence-based medicine for health and care. <https://www.nice.org.uk/guidance/ta254>. 10. NICE. Evidence-based medicine for health and care. <https://www.nice.org.uk/guidance/ta254>. 11. NICE. Evidence-based medicine for health and care. <https://www.nice.org.uk/guidance/ta254>.

Disclosures

All authors are employees of PPD™ Evidera™ Health Economics & Market Access, Thermo Fisher Scientific. This poster was funded by Thermo Fisher Scientific. Editorial and graphic design support were provided by Katerina Calabuig and Richard Leeson of Thermo Fisher Scientific.



Poster Title

Comparing Alternative Extrapolation Methods Using Standard Partitioned Survival Model Functionality in the Presence of Converging Survival Data: A Case Study in Renal Cell Carcinoma

Objective

- The objective of this case study was to explore the impact of combining typical functionalities (treatment effect waning, KM+parametric fit) included in partitioned survival models (PSMs) with standard parametric models as alternative extrapolation approaches, compared to conservatively assuming equivalence at the crossing point or unadjusted extrapolation, considering multiple OS data cuts from the CLEAR trial.

AUTHORS

Paulina Kazmierska¹
Neda Aminnejad²
George Bungey¹

AFFILIATION

¹Thermo Fisher Scientific, London, UK
²Thermo Fisher Scientific, Toronto, Canada

PRACTICE AREA

PPD™ Evidera™ Health Economics & Market Access, Thermo Fisher Scientific

CLIENT

INTERNAL

Comparing Alternative Extrapolation Methods Using Standard Partitioned Survival Model Functionality in the Presence of Converging Survival Data: A Case Study in Renal Cell Carcinoma

Paulina Kazmierska,¹ Neda Aminnejad,² George Bungey¹

¹Thermo Fisher Scientific, London, UK; ²Thermo Fisher Scientific, Toronto, Canada

Thermo Fisher
SCIENTIFIC

EE141

Background

- In specific circumstances, combinations of parametric survival extrapolations may produce implausible crossings between two comparators due to the characteristics of the underlying Kaplan-Meier (KM) data.
- This was observed with the CLEAR trial data in NICE TA858⁵ when comparing lenalidomide in combination with pembrolizumab (LEN+PEM) versus sunitinib (SUN) in patients with renal cell carcinoma (RCC).
- In the August 2020 cut of CLEAR trial data, while LEN+PEM showed improved overall survival (OS) in the short term, crossing occurred between the KM curves at approximately 33 months.

Objectives

- The objective of this case study was to explore the impact of combining typical functionalities (treatment effect waning, KM+parametric fit) included in partitioned survival models (PSMs) with standard parametric models as alternative extrapolation approaches, compared to conservatively assuming equivalence at the crossing point or unadjusted extrapolation, considering multiple OS data cuts from the CLEAR trial.

Methods

- Published OS data from the CLEAR trial were digitized using WebPlotDigitizer software² and pseudo-individual patient data generated by Guyot Algorithm.³ A series of standard parametric distributions was then fitted in the R flexsurv package to the August 2020 data cut.
- Six standard parametric functions were then fitted (exponential, Weibull, lognormal, loglogistic, Gompertz, generalized gamma). Models were selected based on statistical, visual fit, comparisons of hazard profiles for the parametric models against smoothed hazard plots for the KM data, and on the basis of plausibility of long-term predictions according to UK clinical expert expectations for SUN from technology appraisal (TA) 858 by the National Institute for Health and Care Excellence (NICE; <20% at 10 years).⁴ In line with NICE decision support unit (DSU) technical support document (TSD) 14 guidance,⁵ the same type of parametric model was selected for both comparators in the absence of a strong rationale to support different types.
- Given the crossing of OS KM curves observed in the CLEAR trial and potential uncertainty around long-term extrapolations for OS, the following approaches were also explored: (1) assuming equivalence at the crossing point, (2) assuming equivalent efficacy to SUN for LEN+PEM at the start of convergence between the two treatments, after which LEN+PEM OS hazards are set equal to SUN; (3) exploring a combined KM+parametric extrapolation approach using a truncated KM curve.
- Long-term extrapolations were then visually compared with the final data cut KM curve (July 2022).⁶

Results

Unadjusted Extrapolations

- Joint parametric distributions were not suitable due to curve crossings on the log-cumulative hazard plot with clearly non-parallel hazard plots, as well as the result of the formal assessment of the proportional hazards (PH) assumption via the Schoenfeld residuals test (P<0.0001).
- Among individual fits, the Gompertz and log-normal distributions produced the best statistical fit according to both Akaike information criterion (AIC) and Bayesian information criterion (BIC) for LEN+PEM and SUN, respectively. Most SUN fits showed relatively poor visual fit to the observed data, and almost all produced aggressive curve crossing (Figure 1 and Figure 2). The exponential models, while producing poor visual fits, were considered the most plausible set of single-fit distributions given clinical expert expectations for long-term OS discussed in NICE TA858⁵ (<20% for patients starting treatment with SUN) without curve crossing and were selected for both comparators in line with NICE DSU TSD14 guidance⁵.
- Among the remaining individual fits, the Weibull model produced the least sharp curve crossing and met clinical expert expectations from NICE TA858⁵ for long-term OS in the SUN arm. This model was explored for both comparators in the equivalence assumption approach.

Figure 1. Long-term Single Parametric OS Predictions for LEN+PEM Using August 2020 Data Cut

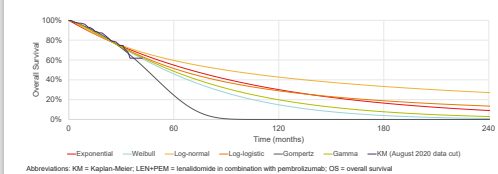
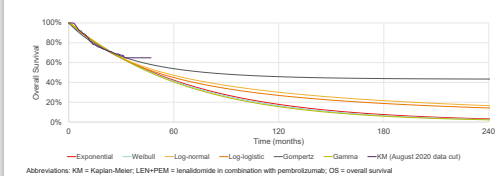


Figure 2. Long-term Single Parametric OS Predictions for SUN Using August 2020 Data Cut

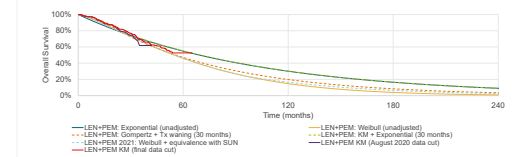


Results (cont.)

Method 1: Equivalence Assumption

- Single Weibull models were applied in combination with the SUN equivalency assumption mechanic to allow for a smoother convergence between LEN+PEM and SUN curves without actual crossing. This approach helped avoid overpredicting the tail of the original KM curve; however, it resulted in more conservative extrapolation for LEN+PEM compared with the unadjusted individual exponential fit and underprediction of the tail of the final KM curve.

Figure 3. Approaches for Long-term OS Predictions for LEN+PEM



Abbreviations: KM = Kaplan-Meier; LEN+PEM = lenalidomide in combination with pembrolizumab; OS = overall survival; SUN = sunitinib; Tx = treatment in Methods 2 and 3; two switch points were explored (24 and 30 weeks); as curves are closely overlapping, only 30-week switch point data is visualized for both methods.

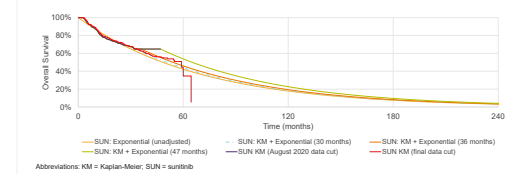
Method 2: Treatment Effect Waning

- The treatment effect waning approach for LEN+PEM explored two switch points: (1) at 24 and (2) at 30 weeks, after which the OS hazard of LEN+PEM was set equal to that of SUN. The switch points were selected based on visual inspection of the KM curves; at these times, curves appeared to begin converging. For LEN+PEM the Gompertz model, the best statistical fit according to both AIC and BIC, was used.
- Both switch points produced nearly overlapping sets of curves. In both cases, the treatment effect waning approach helped avoid overpredicting the tail of the original KM curve, similar to Method 1, while producing a conservative extrapolation for LEN+PEM compared with the unadjusted individual exponential fit with underprediction of the tail of the final curve.

Method 3: KM+Parametric Extrapolation

- The KM+parametric extrapolation approach for LEN+PEM used an individual exponential fit and explored two switch points: (1) at 24 and (2) at 30 weeks; at these times, curves appeared to begin converging on the original KM plot. Both switch points produced nearly overlapping sets of curves. Both approaches resulted in a slightly better visual fit to the observed data to the original unadjusted exponential fit. Compared with the KM curves, the KM+parametric extrapolation approach produced an overprediction of the original 2020 data cut KM data but produced a close fit to the tail of the final 2022 data cut KM curve.
- The KM+parametric extrapolation approach for SUN used an individual exponential fit and explored three switch points: (1) at 47 months, corresponding to the maximum follow-up for the SUN arm at the August 2020 data cut; (2) at 36 months, when the number of patients at risk dropped below 10; and (3) at 30 months, shortly before the original curves crossed. Switching at 47 months resulted in a clinically implausible estimate (>20% at 10 years), whereas the other two switch points resulted in slightly more optimistic extrapolations compared to the original unadjusted exponential fit.

Figure 4. Approaches for Long-term OS Predictions for SUN



Abbreviations: KM = Kaplan-Meier; SUN = sunitinib.

Conclusions

- The results of our case study show that relatively straightforward functionalities commonly implemented in PSMs may provide reasonable alternative extrapolations to support further scenario analyses in the presence of clinically implausible converging survival extrapolations between study arms, especially in cases where proportional hazards are clearly violated and joint parametric fits appear inappropriate, as seen for the CLEAR trial data.
- However, careful rationalization of appropriate switch points for treatment effect waning and KM+parametric approaches is required.

References

- Motzer RJ, et al. *N Engl J Med*. 2021;384(14):1289-300.
- WebPlotDigitizer, version 5.2. <https://automer.io>
- Guay F, et al. *BMC Med Res Methodol*. 2012;12:9. <https://www.ncbi.nlm.nih.gov/pubmed/22525076>
- National Institute for Health and Care Excellence (NICE). TA858: Lenalidomide with pembrolizumab for untreated advanced renal cell carcinoma. 2023. <https://www.nice.org.uk/guidance/ta858>
- Letimer NR. NICE DSU Technical Support Document 14: Understanding survival analysis for economic evaluation alongside clinical trial: extrapolation with patient-level data. 2011. <https://www.ncbi.nlm.nih.gov/pubmed/22525076>
- Motzer RJ, et al. *J Clin Oncol*. 2024;42(11):1222-8.

Disclosures

All authors are employees of PPD™ Evidera™ Health Economics & Market Access, Thermo Fisher Scientific. The views expressed in this study are those of the authors and not necessarily those of Thermo Fisher Scientific. The authors do not have any conflicts of interest.

Acknowledgments

The authors would like to thank Caroline Cole and Kawthar Nakayama of Thermo Fisher Scientific for their editorial and graphic contributions.

Poster Title

Are Clinicians Ready For A More Environmental-friendly Health Care? A Clinician Survey

Objective

- To identify examples for how environmental sustainability has influenced clinical practices to date and to understand expectations for its future impact.
- To test attitudes towards considering sustainability in healthcare decisions and to measure conceptual acceptance of trading off efficacy and safety outcomes, costs, and physician time for environmental benefit.

<p>AUTHORS Kinga Marczell¹ Ruth Chapman² Michael J Cohen³ Ágnes Benedict⁴</p>	<p>AFFILIATION ¹Thermo Fisher Scientific, Budapest, Hungary ²Thermo Fisher Scientific, London, UK ³Thermo Fisher Scientific, Waltham, MA, USA ⁴Thermo Fisher Scientific, Vienna, Austria</p>	<p>PRACTICE AREA PPD™ Evidera™ Health Economics & Market Access, Thermo Fisher Scientific</p>	<p>CLIENT INTERNAL</p>
--	---	--	-----------------------------------

Are Clinicians Ready For A More Environmental-friendly Health Care? A Clinician Survey



HSD119

Kinga Marczell,¹ Ruth Chapman,² Michael J Cohen,³ Ágnes Benedict⁴
¹ Thermo Fisher Scientific, Budapest, Hungary; ² Thermo Fisher Scientific, London, UK; ³ Thermo Fisher Scientific, Waltham, MA, USA; ⁴ Thermo Fisher Scientific, Vienna, Austria

Introduction

- Health care's environmental impact exceeds that of the airline industry,¹ prompting growing interest in addressing this issue within the industry and HTA bodies.
- Considering environmental impact often involves trade-offs, despite assertions to the contrary.
- To promote sustainability in medical practice, it's essential to shape clinical practitioners' attitudes and build consensus on acceptable trade-offs.

Objectives

- To identify examples for how environmental sustainability has influenced clinical practices to date and to understand expectations for its future impact.
- To test attitudes towards considering sustainability in healthcare decisions and to measure conceptual acceptance of trading off efficacy and safety outcomes, costs, and physician time for environmental benefit.

Methods

- A questionnaire was sent to Thermo Fisher Scientific physicians in June 2025.
- It covered changes seen, proposed or foreseen, personal attitudes towards sustainability, and expectations about future changes.
- Multiple-choice and open-ended questions aimed to understand considerations and barriers to embracing changes. A projective question was included to gain insight into genuine attitudes towards environmental outcomes in clinical decisions.
- To explore attitude heterogeneities, questions about environmental consciousness in private life (waste recycling, travel, consumption) were linked to professional decisions.
- Questions also explored willingness to trade-off health outcomes, costs, and physician time for environmental benefits.

Results

Participant characteristics

- Out of 67 participants, 29 responses were received: 38% male; 70% aged 45-64, and 30% aged 35-44. 79% were from the USA, with others from Europe (13%), UK (3%), and Africa (3%).

- 12 practice areas were represented (oncology, allergology, immunology, geriatrics, emergency medicine, family medicine, respiratory, psychiatry, obesity medicine, diabetology, internal medicine, research medicine), and 63% of respondents are active clinicians.

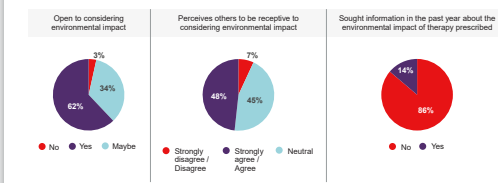
Perception of landscape

- 45% noticed changes in clinical practices for sustainability in the past 5 years, but only 21% could provide specific examples. Only one participant (3%) cited examples involving pharmaceutical selection due to sustainability.
- Observed changes included using asthma inhalers with lower carbon emissions, anaesthesiology gases with reduced global warming potential, decreased printer paper and single-use items, and improved waste management.
- Foreseen changes included reducing toxic materials in pharmaceutical production, fewer single-use plastics, a shift towards virtual trials, and improved medication packaging.

Awareness of and openness to considering environmental information

- Most participants (52%) considered themselves somewhat familiar with sustainability in healthcare; 7% were very familiar, 31% were not very familiar, and 10% were not at all familiar.
- 62% were open to considering environmental outcomes in clinical decisions, but only 48% believed their peers would do the same. Additionally, only 14% actively sought information about the environmental impact of materials or therapies in the past year.

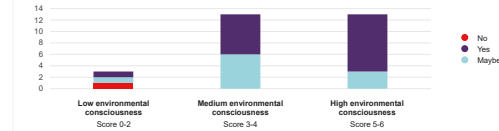
Figure 1. Willingness to consider environmental outcomes in clinical decisions



- Willingness to consider environmental information in decision-making did not vary significantly by gender or age.
- Seeking and considering information professionally were positively correlated with concern for these factors in personal decisions (consumption choices, recycling, travel).

Results (cont.)

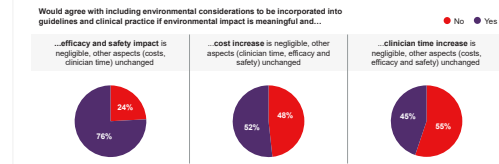
Figure 2. Heterogeneity in the willingness to consider environmental outcomes in clinical decisions



Score on considering environmental aspects in personal decision making is calculated based on responses to three questions: Do you take into account sustainability in your everyday life (i.e. outside of work) in the following areas: 1) Waste management (e.g. recycling), 2) Travel, and 3) Consumption choices (e.g., food from local sources). "Never" = 0 points, "Sometimes" = 1 point, "Always" = 3 points. Maximum achievable score is 6.

- For meaningful environmental improvement, 45%, 52%, and 76% of respondents were willing to accept a negligible worsening of clinician time, costs, and efficacy and safety outcomes, respectively. This question aimed to measure openness to the concept of environmental trade-off, as it proposes a meaningful benefit on environmental outcomes at the price of a negligible sacrifice in the other dimensions.

Figure 3. Conceptual acceptance of trading off health outcomes, costs, and physician time against environmental benefit



The question aimed to measure openness to the concept of trading off, as it proposes a meaningful benefit on environmental outcomes at the price of a negligible sacrifice on the other dimensions.

Strengths and limitations

- Our survey reached a diverse set of physicians across various practice areas and covered multiple dimensions of environmental impact: perception of the current landscape, expectations for future changes, awareness, willingness to consider environmental impact information, and attitudes towards trade-offs.
- There are important limitations to this exploratory survey. The sample is not representative of all physicians due to multiple factors. Physicians employed by Thermo Fisher Scientific may have different attitudes compared to other physicians, and the direction and magnitude of this difference are unknown. Results are subject to selection bias; those with a favorable attitude towards environmental impacts may have been more likely to respond.

Conclusions

- Most clinicians observed changes in clinical practices aimed at improving sustainability over the past five years, and many identified areas for future changes.
- Clinicians showed a cautiously positive attitude towards considering environmental aspects in healthcare decisions, but there was significant heterogeneity in awareness and stance. They were divided on willingness to trade off efficacy, safety outcomes, costs, and physician time for environmental benefits. Many were unwilling to make even negligible sacrifices for meaningful environmental improvements, especially regarding their own time.
- Overall, while there is considerable awareness and openness among clinicians towards integrating environmental considerations into healthcare decisions, diverse attitudes towards accepting trade-offs highlight the need for further education and consensus-building to achieve meaningful progress in sustainable healthcare practices.

References

- Berners-Lee, M. *The Carbon Footprint of Everything*. Greystone Books; 2022.

Disclosures

MM and RC are employees of PPD™ Evidera™ Health Economics & Market Access, Thermo Fisher Scientific. MJC is an employee of Thermo Fisher Scientific. AB was an employee of Thermo Fisher Scientific at the time of conducting the study. This poster was funded by Thermo Fisher Scientific.

Acknowledgments

The authors would like to thank Craig Boyle, DO, FACEP, TrialMed part of Thermo Fisher Scientific for his support of this project and helping physicians across Thermo Fisher Scientific. Editorial and graphic design support were provided by Caroline Cole and Sarah Berger of Thermo Fisher Scientific.



Poster Title

Navigating Health Technology Assessment Requirements: The Current Landscape of Alzheimer's Disease Modifying Treatments

Objective

- Our research aims to provide an overview of DMT health technology assessment (HTA) appraisals and highlight the challenges faced by manufacturers when the clinical value and cost-effectiveness of DMTs for AD are under evaluation.

AUTHORS Victoria Molenkamp ¹ Nathalie Abi-Saleh ² Denise So ¹	AFFILIATION ¹ Thermo Fisher Scientific, London, UK ² Thermo Fisher Scientific, Waltham, MA, USA	PRACTICE AREA PPD™ Evidera™ Health Economics & Market Access, Thermo Fisher Scientific	CLIENT INTERNAL
--	--	--	---------------------------



Navigating Health Technology Assessment Requirements: The Current Landscape of Alzheimer's Disease Modifying Treatments

Thermo Fisher SCIENTIFIC

HTA244

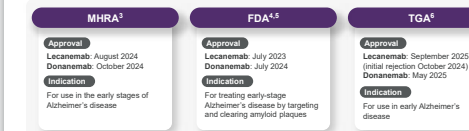
Victoria Molenkamp,¹ Nathalie Abi-Saleh,² Denise So¹

¹Thermo Fisher Scientific, London, UK; ²Thermo Fisher Scientific, Waltham, MA, USA

Introduction

- The prevalence of dementia worldwide is projected to reach 78 million cases by 2030.¹
- Alzheimer's disease (AD) is the most common form of dementia and accounts for up to 70% of all cases.²
- As of September 2025 (Figure 1), two disease-modifying treatments (DMTs) for AD have received authorisations in Europe, the US, and Australia: lecanemab (Leqembi) and donanemab (Kinsula).

Figure 1. Regulatory approvals of DMTs across the MHRA, FDA and TGA



Abbreviations: DMT = disease modifying treatment; FDA = Food and Drug Administration; MHRA = Medicines and Healthcare Products Regulatory Agency; TGA = Therapeutic Goods Administration

Objectives

- Our research aims to provide an overview of DMT health technology assessment (HTA) appraisals and highlight the challenges faced by manufacturers when the clinical value and cost-effectiveness of DMTs for AD are under evaluation.

Methods

- We conducted a targeted search for published materials on the assessment of DMTs in AD, including HTA appraisals, press releases, and reports in the EU4, the UK, Australia, Canada, and the US.
- Searches were run on 18 September 2025 across nine HTA body websites.

Results

- The results of the searches are presented in Figure 2.
- We identified nine appraisals in total (two for National Institute for Health and Care Excellence [NICE], two for Scottish Medicines Consortium [SMC], two for Institute for Clinical and Economic Review [ICER], and one each for Gemeinsamer Bundesausschuss [G-BA], Haute Autorité de Santé [HAS], and the Pharmaceutical Benefits Advisory Committee [PBAC]).
- We did not identify any published DMT appraisals for AD in Italy, Spain, or Canada.
- Our search identified one NICE Innovation Laboratory report that highlighted potential challenges faced by manufacturers when DMTs are assessed for dementia.³
- Key uncertainties included prevalence estimates for the eligible treatment population, long-term treatment benefit, whether trial outcomes are clinically meaningful, and the additional cost of service implications.
- The report concluded that while multiple challenges may arise during the evaluation of DMTs, NICE's approach is appropriate for their assessment.

Figure 2. Overview of relevant publications identified in the TLR



Abbreviations: AEMPS = Agencia Española de Medicamentos y Productos Sanitarios; AIFA = Agenzia Italiana del Farmaco; CDA-AMC = Canada's Drug Agency; G-BA = Gemeinsamer Bundesausschuss; HAS = Haute Autorité de Santé; ICER = Institute for Clinical and Economic Review; NICE = National Institute for Health and Care Excellence; PBAC = Pharmaceutical Benefits Advisory Committee; SMC = Scottish Medicines Consortium; TLR = targeted literature search

Results (cont.)

- An overview of HTA recommendations is presented in Table 1.

Table 1. Overview of HTA recommendations for DMTs in the treatment of AD

Negative recommendations	Positive recommendations
<ul style="list-style-type: none">• England and Scotland: Neither donanemab nor lecanemab are recommended for use following appraisals conducted between May and July 2025. Both are due for appeal with NICE in October 2025.⁷ Reasons cited by NICE and SMC for the negative recommendations included limited evidence to demonstrate long-term effects, uncertain cost-effectiveness, and minimal benefits relative to the high costs associated with infusions and monitoring for adverse events.^{7,8} The SMC specifically queried whether the benefit with donanemab was clinically meaningful.⁸• France: HAS rejected the request for early-access use of lecanemab for AD. The decision cited a modest level of efficacy deemed not clinically meaningful, a concerning safety profile, the need for frequent MRI monitoring, and a lack of a robust QoL assessment.⁹• Australia: PBAC rejected the PBS application of donanemab, noting a high burden for patients and the health system, substantial risks, modest clinical impact, uncertainty that trial results would translate into meaningful improvements, and the requirement for frequent infusions and monitoring.¹⁰• US: ICER determined that currently available evidence is insufficient to demonstrate a net benefit of lecanemab over BSC. It concluded that the safety risks and high list price of lecanemab may outweigh its efficacy benefits.¹¹	<ul style="list-style-type: none">• Germany: Lecanemab received funding as the first monoclonal antibody available for early AD treatment in Germany because it offers a new DMT option in this indication. However, the scope has been restricted due to safety risks and high diagnostic/monitoring demands. Therefore, use is restricted to selected patients under specialist supervision, with mandatory checks every 6 months and controlled access programme registration to monitor patient progress.¹²• US: ICER did not assess donanemab because it received accelerated FDA approval.¹¹ Both lecanemab and donanemab are reimbursed by Medicare.¹³

Abbreviations: AD = Alzheimer's disease; BSC = best supportive care; DMT = disease modifying treatment; HAS = Haute Autorité de Santé; HTA = health technology assessment; ICER = Institute for Clinical and Economic Review; NICE = National Institute for Health and Care Excellence; PBAC = Pharmaceutical Benefits Advisory Committee; PBS = Pharmaceutical Benefits Scheme; QoL = quality of life; SMC = Scottish Medicines Consortium

Considerations for the future

- DMTs present unique challenges to HTA due to uncertainties related to long-term effectiveness, cost-effectiveness, and clinically meaningful patient benefit. The majority of HTA assessments for donanemab and lecanemab have resulted in negative recommendations.
- As of October 2025, NICE is the only HTA body that has published a review of challenges and considerations in the evaluation of DMTs. There are currently three DMT appraisals in development (hydromethylthionine mesylate) or awaiting development (gantenerumab and bapinecesumab) for NICE, which may provide additional insights once published. However, there is a need for additional guidance that captures the HTA requirements and priorities across different markets.

Conclusions

- Our findings highlight a potential disconnect between HTA evidence requirements for DMTs in AD and the available evidence provided in appraisals. This is particularly relevant given that three additional appraisals for DMTs in AD are awaiting development at NICE at the time of writing and more appraisals are expected in the future. Findings from our research may also be applicable to other DMTs for neurodegenerative diseases, such as pridopidine for treating Huntington's disease.

References

1. Alzheimer Disease International. Dementia statistics. 2025. <https://www.alzint.org/about/dementia-facts-figures/dementia-statistics>.
2. World Health Organization. Dementia. 2025. <https://www.who.int/news-room/fact-sheets/detail/dementia>.
3. Medicines & Healthcare products Regulatory Agency. <https://www.gov.uk/government/organisations/medicines-and-healthcare-products-regulatory-agency>.
4. Biogen. FDA approves LECANEMAB (Leqembi™) (lecanemab-imb) intravenous injection for maintenance dosing for the treatment of early Alzheimer's disease. <https://investor.biogen.com/news-releases/news-release-details/fda-approves-leqembi-qilkm-lecanemab-imb-intravenous>.
5. Lilly. Lilly's Kinsula™ (donanemab-ahbl) approved by the FDA for the treatment of early symptomatic Alzheimer's disease. <https://investor.lilly.com/news-releases/news-release-details/lillys-kinsula-donanemab-ahbl-approved-fda-treatment-early>.
6. Therapeutic Goods Administration. TGA confirms decision to not register lecanemab (LEQEMBI). 2025. <https://www.tga.gov.au/news/news/tga-confirms-decision-not-register-lecanemab-leqembi>.
7. National Institute for Health and Care Excellence. <https://www.nice.org.uk/guidance/TA988>.
8. Scottish Medicines Consortium. <https://scottishmedicines.org.uk/>.
9. Fondation Alzheimer Leqembi (Lecanemab). La Haute Autorité de Santé rejette la demande d'accès précoce. <https://www.fondation-alzheimer.org/leqembi-lecanemab-la-haute-autorite-de-sante-rejette-la-demande-dacces-precoce>.
10. New Alzheimer's drug deemed 'unsuitable' for PBS. 2025. <https://www7.rtrg.org.au/news/2025/alzheimer-s-drug-deemed-unsuitable-for-pbs>.
11. Institute for Clinical and Economic Review. What is ICER? <https://icer.org/what-is-icer/>.
12. Gemeinsamer Bundesausschuss. <https://www.g-ba.de/>.
13. Statement: broader Medicare coverage targets available following FDA traditional approval. 2025. <https://www.cms.gov/newsroom/press-releases/statement-broader-medicare-coverage-targets-available-following-fda-traditional-approval>.

Disclosures

All authors are employees of PPD™ Evidera™ Health Economics & Market Access, Thermo Fisher Scientific, which provides consulting and other research services to pharmaceutical, medical device, and related organizations. In their stated positions, they work with a variety of companies and organizations and are precluded from receiving payment or honors directly from these organizations for services rendered. This poster was funded by Thermo Fisher Scientific.

Acknowledgments

German translation support was provided by Dr Pia Gähler of Thermo Fisher Scientific. Editorial and graphic design support were provided by Karissa Calara and Shari Berger of Thermo Fisher Scientific.



Poster Title

A Layered Approach to Reducing Hallucinations in LLMs for Clinical Data

Objective

- The main objective of this study is to reduce hallucinations in LLMs when applied to structured (e.g., tabular EHR data) and unstructured (e.g., clinical text) sources.
- To design a two-layer framework where (1) refined prompts enforce factual, contextspecific outputs and (2) LangChain operationalizes RAG with agentic control.
- To demonstrate, through case examples, how prompt validation and retrieval-aware orchestration improve factual grounding and transparency in clinical and research workflows.
- To provide a scalable, interpretable, and compliant approach for deploying LLMs in healthcare analytics and decision support.

AUTHORS Ashwin Rai ¹ Devika Bhandary ² Victoria Ikoro ³ Andre Ng ³	AFFILIATION ¹ Thermo Fisher Scientific, Waltham, MA, USA ² Thermo Fisher Scientific, London, UK ³ Thermo Fisher Scientific, Ontario, Canada	PRACTICE AREA PPD™ Evidera™ Real-World Data & Scientific Solutions, Thermo Fisher Scientific	CLIENT INTERNAL
---	--	--	---------------------------





MSR4

A Layered Approach to Reducing Hallucinations in LLMs for Clinical Data

Ashwin Rai,¹ Devika Bhandary,² Victoria Ikoro,³ Andre Ng³
¹Thermo Fisher Scientific, Waltham, MA, USA; ²Thermo Fisher Scientific, London, UK; ³Thermo Fisher Scientific, Ontario, Canada

Introduction

- Large language models (LLMs) are being piloted to automate literature review, summarize electronic health records (EHRs), and extract clinical insights from structured and unstructured data.^{1,2}
- Their promise is speed and scale, but reliability is critical when outputs inform care or policy.
- Hallucinations, defined as fabricated or misleading content, remain a key risk without careful prompt design and operational safeguards.³
- These risks are heightened in healthcare, where inaccurate or nontransparent outputs may compromise trust and decision-making.^{2,3}

Problem statement

- Naive prompting (single "stuff" chain): With a generic "answer based on context" template and no retrieval guardrails, the LLM produced speculative outputs and failed to tie claims to evidence, such as vague diabetes summaries and unsupported hemoglobin A1C (HbA1c) generalizations when specific data were required.
- This behavior illustrates hallucination risk and misuse of context in clinical questions and answers (QA), underscoring the need for structured controls.³

Proposed remedy (layered RAG and agent)

- Layer 1 – Prompting & Validation:** Domain-specific, context-only prompts, which enforce factuality and signal when evidence is absent.⁴
- Layer 2 – Orchestration (LangChain):** Modular chains with prompt templates, retrieval-augmented generation (RAG),^{4,5} and lightweight agentic control⁶ to separate instructions from logic, route to appropriate models, and ensure only retrieved evidence is used.

Objectives

- The main objective of this study is to reduce hallucinations in LLMs when applied to structured (e.g., tabular EHR data) and unstructured (e.g., clinical text) sources.
- To design a two-layer framework where (1) refined prompts enforce factual, context-specific outputs and (2) LangChain operationalizes RAG with agentic control.
- To demonstrate, through case examples, how prompt validation and retrieval-aware orchestration improve factual grounding and transparency in clinical and research workflows.
- To provide a scalable, interpretable, and compliant approach for deploying LLMs in healthcare analytics and decision support.

Methods (cont.)

Figure 2. Steps for the proposed methodology framework

Abbreviations: LLM = large language model; QA = questions and answers; RAG = retrieval-augmented generation

Results

- The framework was tested across different input sources, including structured EHR-like data and unstructured clinical text.
- As shown in Table 1, naive prompting produced generalized or speculative outputs, whereas the layered RAG and agent approach generated precise, evidence-grounded responses.

Table 1. Illustration of prompt and response showing naive prompting (generic) versus layered RAG and agent (grounded)

Aspect	Naive Prompting (Baseline)	Layered RAG and Agent Framework
Transparency and accuracy	Generic, speculative, low accuracy (e.g., "Diabetic often have high HbA1c")	Grounded, specific, high accuracy (e.g., "HbA1c = 8.1% (diabetic) vs. 5.5% (non-diabetic)")
Evidence use	No link to retrieved data; outputs not tied to context	Strictly based on retrieved context; answers reference-extracted rows
Clinical reliability	Risk of misinformation; low trustworthiness; unreported summaries	Traceable, reliable, aligned with domain requirements; available outputs
Reproducibility	Variable outputs; little consistency; different answers to same prompt	Consistent workflows through modular orchestration; structured chain ensures repeatability
Context handling	Weak; often ignores or misuses data; skips available demographics	Strong; explicitly tied to structured and unstructured inputs; finds highest HbA1c segment
Output interpretability	Little clarity on reasoning or evidence path; no way to verify sources	Transparent chain of reasoning; source-linked answers; retrieval and computation shown
Practical utility	Limited for healthcare decision support; not suitable for risk stratification	Suitable for analytics, evidence synthesis, and policy use; supports clinical alignment

Abbreviations: HbA1c = hemoglobin A1c; RAG = retrieval-augmented generation

Figure 3. Illustration of prompt and response showing naive prompting (generic) versus layered RAG and agent (grounded)

Conclusions

- A layered RAG approach substantially reduced hallucinations and improves factual accuracy compared with naive prompting.
- LangChain's modular orchestration framework enables scalability, flexibility, and transparency, supporting reliable deployment in diverse healthcare contexts.
- Together, these methods produce transparent, reproducible, and clinically aligned pipelines that strengthen trust in LLM-assisted healthcare analytics and decision support.
- This approach provides a scalable pathway for adoption, balancing performance with compliance, and paving the way for safer, evidence-based use of LLMs in real-world practice.

References

1. Lewis JH, et al. Retrieval-augmented generation for knowledge-intensive NLP tasks. 2020. 2. Alkhalil B, et al. (Eds) RAG: Learning to retrieve, generate, and critique through self-reflection. 2024. 3. Yao S, et al. ReAct: Prompting reasoning and acting in language models. 2024. 4. LangChain Devs. Prompt Templates (ChatPromptTemplate & ChatPromptTemplate). <https://docs.langchain.com/docs/guides/prompts/templates>. 5. LangChain Devs. Tools (Retrieval-Augmented Generation (RAG) Support). <https://docs.langchain.com/docs/guides/tools/retrieval-augmented-generation>. 6. Duval M, et al. arXiv preprint arXiv:2410.01011. 2024. 7. LangChain Devs. create_react_agent_prompt() & index on https://github.com/langchain-ai/langchain/blob/main/libs/langchain/src/agents/react/output_parser.py. 8. LangChain Devs. AgentExecutor (structured_output_executor). <https://github.com/langchain-ai/langchain/blob/main/libs/langchain/src/agents/executor.py>.

Acknowledgements

Editorial and graphic design support was provided by Caroline Cole and Karissa Calver of Thermo Fisher Scientific.

Presented at ISPOR —The Professional Society for Health Economics and Outcomes Research—Europe | 9–12 November 2025 | Glasgow, Scotland, UK

© 2025 Thermo Fisher Scientific Inc. All rights reserved. All trademarks are the property of Thermo Fisher Scientific and its subsidiaries unless otherwise specified. This information is not intended to encourage use of these products in any manner that might infringe the intellectual property rights of others.




Poster Title

Identifying Risk Profiles for Early Treatment Discontinuation in Geographic Atrophy Using Machine Learning and SHAP Clustering

Objective

- The main objective of this study was to develop a machine learning algorithm to identify patients at risk of dropout.
- For each patient, the algorithm would produce a risk score for adherence based on their individual patient journey.
- By utilizing a data-driven approach, we sought to explore the clinical, behavioral, and psychosocial factors behind patients' decisions to be non-adherent.
- The ultimate goal was to subsequently create a toolkit for healthcare professionals to implement a tailored retention strategy.

<p>AUTHORS</p> <p>Ashwin Rai¹ Ariel Berger¹ Andre Ng²</p>	<p>AFFILIATION</p> <p>Victoria Ikoro³ Devika Bhandary²</p>	<p>PRACTICE AREA</p> <p>PPD™ Evidera™ Real-World Data & Scientific Solutions, Thermo Fisher Scientific</p>	<p>CLIENT</p> <p>INTERNAL</p>
---	---	---	--------------------------------------



MSR126

Identifying Risk Profiles for Early Treatment Discontinuation in Geographic Atrophy Using Machine Learning and SHAP Clustering

Ashwin Rai,¹ Ariel Berger,¹ Andre Ng,² Victoria Ikoro,³ Devika Bhandary²
¹Thermo Fisher Scientific, Waltham, MA, USA; ²Thermo Fisher Scientific, London, UK; ³Thermo Fisher Scientific, Ontario, Canada

Introduction

Geographic atrophy

- Geographic atrophy (GA) is an advanced form of age-related macular degeneration.
- Approximately 1 million US adults are affected in at least one eye.¹
- Risk increases with age (~0.3% of persons aged 65–74 years; ~4% of persons aged ≥ 85 years).²
- It is a progressive disease that results in irreversible loss of vision over time.³

Treatment

- Frequency of administration depends on the drug and regimen selected, but the current mainstay of treatment is intravitreal injections directly into the eye every 25–60 days.⁴
- The main goal of this treatment is to delay GA progression.
- While effective, studies have shown 30%–40% of patients discontinue treatment within a year.⁵

Problem

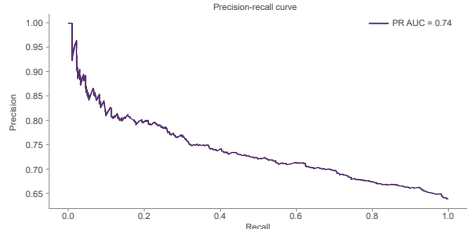
- As treatment is only effective if maintained, relatively high rates of discontinuation reduce patient long-term quality of life.⁶
- However, healthcare professionals do not typically have the time or resources to intensively monitor their patients, and thus support their patients' decisions to adhere to treatment.⁵
- Early discontinuation also affects the accuracy and reliability of clinical studies of treatment for GA.⁷
- Therefore, decision rules/algorithms targeting patient profiles at elevated risk of discontinuation may represent an efficient means to improve adherence and subsequently optimize positive outcomes.
- Specifically, if these algorithms were built on electronic, real-world, health data, they could be automated at various points-of-care to maximize their impact on treatment adherence.

Results

ML model results on patients at risk of GA treatment dropout (Figure 3)

- Prediction target was discontinuation within 120 days.
- The receiver operating characteristic (ROC) area under the curve (AUC) was 0.62; the precision-recall AUC was 0.74, with recall of 0.98 and precision of 0.64
- Overall accuracy was 64%.

Figure 3. Precision-recall curve by supervised machine learning model identifying patients at risk of GA treatment dropout



Abbreviations: PR AUC= precision-recall area under the curve

Patient profiles comparison (Figure 4)

Profile 3—Lower GA treatment dropout

- 56% of patients in this cluster (which represents 26% of the total population) discontinued treatment within 120 days.
- They had relatively better initial vision, stable healthcare access, less severe disease (bilateral GA/fovea patterns), slower disease progression, and most were treatment naive (i.e., first GA treatment ever).

Profile 5—Higher GA treatment dropout

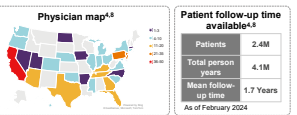
- 75% of patients in this cluster (which represents 14% of the total population) discontinued treatment within 120 days.
- They had relatively poorer initial vision, limited/inadequate healthcare coverage, among other disease traits.

Objectives

- The main objective of this study was to develop a machine learning algorithm to identify patients at risk of dropout.
- For each patient, the algorithm would produce a risk score for adherence based on their individual patient journey.
- By utilizing a data-driven approach, we sought to explore the clinical, behavioral, and psychosocial factors behind patients' decisions to be non-adherent.
- The ultimate goal was to subsequently create a toolkit for healthcare professionals to implement a tailored retention strategy.

Methods

Data source – Vestrum retina database

<p>Database overview^{4,5}</p> <p>Patient Visits (2015–Current) 18M Patients (2015–Current) 2.4M Sites (Current) 74 Retina Specialists (Current) 370 States (US) Represented (Current) 35</p>	<p>Physician map^{4,5}</p> 	<p>Patient follow-up time available^{4,5}</p> <p>Patients 2.4M Total person years 4.1M Mean follow-up time 1.7 Years As of February 2024</p>
--	---	---

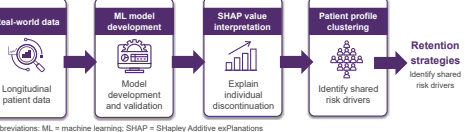
In addition, we can mine the free-text field within the electronic medical record (EMR) to discover further diagnostics and characteristics like subretinal fluid and vitreomacular traction (VMT)

- Demographics:** Age, gender, smoking status, alcohol use
- Treatment history:** Medications/treatment regimens (prior and present) and related procedures
- Disease characteristics:** Relevant disease (e.g., GA) and comorbidities, year and severity of diagnosis, disease progression, adverse events
- Patient-reported outcomes:** Visual acuity, central retinal thickness, fluid, surgery outcomes, change in therapy, lost to follow-up
- Targeted adverse events:** infection (uveitis, endophthalmitis), retinal detachments, hemorrhaging, disease development, and progression

ML framework (Figure 1)

- We modelled early treatment discontinuation by training a supervised ML model based on 10,000+ GA patients; it was trained on 8,134 patients, tested on 2,034 patients (80–20% split)
- Model inputs included demographics, clinical history, and treatment patterns.
- We used SHapley Additive ExPlanations (SHAP) to interpret patient-level risk drivers and applied K-means clustering on SHAP values to group similar patients. We identified five distinct patient profiles based on shared dropout risk factors.

Figure 1. Overview of risk profiling model and real-world impact



Abbreviations: ML = machine learning; SHAP = SHapley Additive ExPlanations

Conclusions

- Patient profiles translated complex ML outputs and diverse real-world-data into clear, actionable subgroups.
- The profiles revealed both who is at risk or has a specific need, and why, enabling targeted strategies.
- The GA discontinuation example shows how prediction → explanation → segmentation can improve persistence and outcomes.
- Profiling is flexible—applicable to risk, engagement, value, and other use cases across the patient journey.
- Embedding profiles into workflows supports precision interventions, optimizes resources, and strengthens real-world evidence.

References

1. Kim DS, et al. JAMA Ophthalmol. 2023;141(12):1302-8. 2. Yau F, Finkel M. Geographic atrophy. 2017. <https://doi.org/10.1016/j.jco.2017.03.003>. 3. Finkel M, et al. Proc OpticNet (Luscombe). 2023;3:1377803. 4. PPD™ clinical research business of Thermo Fisher Scientific. PPD Vestrum Health™ Real-World Data. <https://www.ppd.com/en/real-world-data>. 5. Bhandary D, et al. Eye. 2023;37(11):152. 6. Nissen ANK, et al. Ophthalmol Ther. 2025;14(8):1173-81. 7. Rosenblatt TR, et al. Graefes Arch Clin Exp Ophthalmol. 2022;20(1):93-100. 8. American Society of Retina Specialists. Retina image bank. <https://retinabank.asrs.org/>

Disclosures

All authors are employees of PPD™ Evidera™ Real-World Data & Scientific Solutions, Thermo Fisher Scientific. This poster was funded by Thermo Fisher Scientific.

Acknowledgment

Editorial and graphic design support were provided by Candice Cole, Karissa Catara, and Richard Leason of Thermo Fisher Scientific.

Presented at ISPOR —The Professional Society for Health Economics and Outcomes Research—Europe | 9–12 November 2025 | Glasgow, Scotland, UK
© 2025 Thermo Fisher Scientific Inc. All rights reserved. All trademarks are the property of Thermo Fisher Scientific and its subsidiaries unless otherwise specified. This information is not intended to encourage use of these products in any manner that might infringe the intellectual property rights of others.




Poster Title

Uncovering Patient Narratives of Opioid Use and Recovery Using Large Language Models for Topic and Emotion Analysis of Social Media

Objective

- To apply LLM-based topic modeling to patient-authored opioid-related discussions from online forums, identifying major themes in lived experience.
- To perform LLM-based emotion analysis across narratives, characterizing emotional burdens such as fear, sadness, hope, and relief.
- To generate scalable, interpretable insights from unstructured patient narratives that can inform tailored interventions, patient support resources, and health policy.

<p>AUTHORS Ashwin Rai¹ Victoria Ikoro² Devika Bhandary³ Andre Ng³</p>	<p>AFFILIATION ¹Thermo Fisher Scientific, Waltham, MA, USA ²Thermo Fisher Scientific, Ontario, Canada ³Thermo Fisher Scientific, London, UK</p>	<p>PRACTICE AREA PPD™ Evidera™ Real-World Data & Scientific Solutions, Thermo Fisher Scientific</p>	<p>CLIENT INTERNAL</p>
--	---	--	-----------------------------------



MSR208

Uncovering Patient Narratives of Opioid Use and Recovery Using Large Language Models for Topic and Emotion Analysis of Social Media

Ashwin Rai,¹ Victoria Ikoro,² Devika Bhandary,³ Andre Ng³
¹ Thermo Fisher Scientific, Waltham, MA, USA; ² Thermo Fisher Scientific, Ontario, Canada; ³ Thermo Fisher Scientific, London, UK

Background

- Opioid use disorder (OUD) imposes substantial clinical and societal burden, and risks evolve with prescription and illicit opioids.^{1,2}
- Understanding patient perspectives is essential for value-based, patient-centered care.
- Social media forums contain large-scale, unfiltered patient narratives about withdrawal, pain management, recovery, and emotional challenges, but the data are unstructured and difficult to analyze at scale.^{3,4}
- Large language models (LLMs) enable scalable extraction of topics and emotions from these narratives, providing real-world evidence to inform interventions and policy.⁵

Objectives

- To apply LLM-based topic modeling to patient-authored opioid-related discussions from online forums, identifying major themes in lived experience.
- To perform LLM-based emotion analysis across narratives, characterizing emotional burdens such as fear, sadness, hope, and relief.
- To generate scalable, interpretable insights from unstructured patient narratives that can inform tailored interventions, patient support resources, and health policy.

Methods

- We analyzed more than 5,000 opioid-related posts and comments from the r/opiates subreddit to identify emotional and thematic patterns in patient narratives.

Data collection

- Source: Reddit's r/opiates forum
- Sample: >5,000 English-language posts and comments
- Inclusion: Texts related to opioid use, withdrawal, recovery, relapse, or support
- Data were de-identified and pre-processed to reduce noise.

Text embedding and topic clustering

- Each post and comment was embedded using a pre-trained LLM to capture semantic content.
- Unsupervised clustering (e.g., density-based) was applied to these embeddings to group semantically similar texts.
- The resulting clusters were manually labeled based on top keywords and representative samples, yielding 20 interpretable topic clusters.

Emotion classification

- We used a transformer-based classifier trained on the GoEmotions dataset to assign emotions to each post and comment.
- The model's native 28-label output was mapped to 10 emotion categories: *sadness, fear, anger, disgust, anxiety, relief, hope, gratitude, joy, neutral*.
- Neutral labels were excluded from further analysis.

Emotion aggregation and topic emotion profiling

- Within each topic cluster, we calculated the share of posts/comments associated with each non-neutral emotion.
- The dominant emotion per topic was defined as the most frequent high-confidence label.
- Emotional profiles were visualized to highlight variation across opioid-related themes.

Results (cont.)

Emotions by topic cluster

- As shown in Figure 2, most opioid-related discussions were dominated by negative emotions, particularly anxiety (35% of all posts) and sadness (20%). Positive tones were less frequent overall, but joy (15%) and gratitude (6%) highlighted the supportive role of online communities.
- Figure 4 illustrates how these emotions varied by topic cluster. Anxiety dominated high-risk themes such as *Overdose & Nalcan Use, Withdrawal & Dosing, and Pill Authenticity & Safety*. Sadness defined grief-related clusters (*Grief & Family Loss, Pain, Grief & Recovery*), reflecting the depth of bereavement and suffering. In contrast, recovery and peer-support clusters (*Recovery Milestones & Hope, Gratitude & Community Praise*) surfaced more positive emotions, with gratitude, joy, and hope balancing the overall negative climate.
- These findings underscore the emotional heterogeneity of patient narratives and identify potential touchpoints for designing patient-centered support interventions in OUD.

Figure 2. Overall emotional distribution

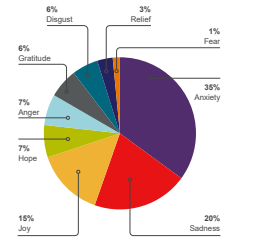
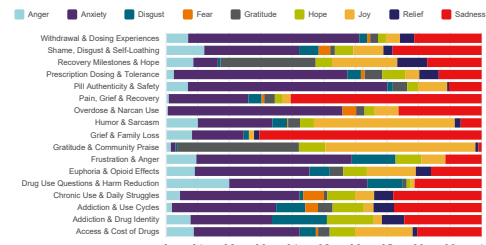


Figure 3. Illustrative sample quotes

- Anxiety / Overdose & Nalcan Use: "I've overdosed twice—now I keep Nalcan next to my bed."
- Sadness / Grief & Family Loss: "I miss my brother. I miss my brother every single day—the pain doesn't go away."
- Gratitude / Recovery Milestones: "90 days clean—thank you all for keeping me strong."

Figure 4. Emotion distribution across 17 topic clusters



Conclusions

- LLM-based clustering and emotion analysis successfully uncovered 17 interpretable topics in opioid-related Reddit discussions, spanning withdrawal, overdose, stigma, recovery, and support.
- Anxiety was the most frequent dominant emotion, especially in high-risk themes such as overdose, pill safety, and withdrawal.
- Sadness was prominent in grief-related topics, while gratitude and joy were most common in recovery and peer-support discussions.
- Emotion–topic mapping from unstructured social data offers a scalable way to surface patient-reported experiences and inform targeted interventions in opioid care.

References

- Yolkow ND, Haney C. *Mol Psychiatry*. 2021;26(11):218–33.
- Nicholls VE, et al. *BMJ open*. 2020;10(2):e002088.
- Sarkar A, et al. *J Med Internet Res*. 2017;19(10):e301.
- Carpenter KA, et al. *PLoS Digital Health*. 2025;4(4):e0000942.
- Thirunavukarasu AJ, et al. *Nat Med*. 2023;29(8):1930–40.

Disclosures


AR, VI, DB and AN are employees of PPD™ Evidera™ Real-World Data & Scientific Solutions, Thermo Fisher Scientific. Funding provided by Thermo Fisher Scientific.

Acknowledgments

Editorial and graphic design support was provided by Caroline Cole and Shari Berger of Thermo Fisher Scientific.

Presented at ISPOR —The Professional Society for Health Economics and Outcomes Research—Europe | 9–12 November 2025 | Glasgow, Scotland, UK

© 2025 Thermo Fisher Scientific Inc. All rights reserved. All trademarks are the property of Thermo Fisher Scientific and its subsidiaries unless otherwise indicated. This information is not intended to encourage use of these products in any manner that might infringe the intellectual property rights of others.



Poster Title

Enhancing Migraine Preference Research: Recommendations from a Systematic Review of Preference Studies

Objective

- To examine how preference studies are designed and reported in the context of migraine treatment, with the aim of informing the design and conduct of future studies.

AUTHORS

Natasha Ramachandran¹
 Harrison Clarke²
 Divya Mohan³
 Jaein Seo²

AFFILIATION

¹Thermo Fisher Scientific, London, UK
²Thermo Fisher Scientific, Waltham, MA, USA
³Patient-Centered Outcomes, Open Health, London, UK

PRACTICE AREA

PPD™ Evidera™ Patient-Centered Research, Thermo Fisher Scientific

CLIENT

INTERNAL
 Patient-Centered Outcomes, Open Health



Enhancing Migraine Preference Research: Recommendations from a Systematic Review of Preference Studies



PCR11

Natasha Ramachandran,¹ Harrison Clarke,² Divya Mohan,³ Jaein Seo²

¹Thermo Fisher Scientific, London, UK; ²Thermo Fisher Scientific, Waltham, MA, USA; ³Patient-Centered Outcomes, Open Health, London, UK

Background

- Migraine management includes acute, preventive, or combined treatments; however, care is often suboptimal due to undertreatment, which presents substantial public health challenges.
- The varying severity, frequency, and characteristics of migraine complicate treatment optimization, requiring individualized plans based on patient preferences.
- Preference studies aim to inform shared decision-making, improve adherence, and guide patient-centered treatment development. However, variations in design, attribute selection, and analysis limit comparability and application. Despite this, no systematic synthesis exists.
- This systematic review synthesized current evidence on treatment preferences in migraine and highlighted valued treatment attributes and methodological patterns.

Methods

- Embase, MEDLINE, and the Cochrane Library were searched for relevant preference studies on migraine treatments; studies must have employed stated-preference methods.
- Two researchers independently screened studies, and disagreements regarding the inclusion or exclusion of studies were resolved through discussion between them; any remaining discrepancies were adjudicated by a third researcher.
- Extracted information included study characteristics, analytical approaches, methods for attribute development, choice task design, and attribute framing.
- Frequencies were calculated by study for study characteristics, analytic approaches, attribute development methods, and choice task design, and by attribute for attribute framing.
- To compare the importance of attribute concepts, rankings or quantitative importance measures (e.g., marginal utilities) were extracted from studies including at least two distinct attribute concepts, excluding contingent valuation methods (CVM) studies. Explicit rankings were used directly; otherwise, attributes were ranked within each study based on the magnitude of reported values.
- Benefit attributes were grouped at the concept level (e.g., speed of onset, durability of effectiveness), while others were synthesized at the sub-concept level (e.g., injection-site reaction, nausea). Average ranks across studies overall indicated overall perceived importance (i.e., the most important attribute ranked=1).

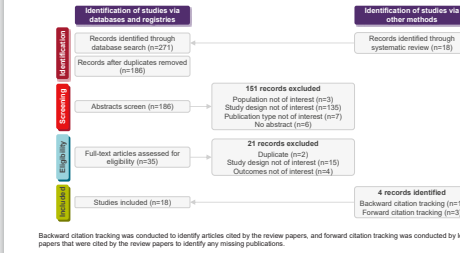
Objective

- To examine how preference studies are designed and reported in the context of migraine treatment, with the aim of informing the design and conduct of future studies.

Results

- Two-hundred and seventy-one studies were identified from the literature search and screened. Eighteen studies were deemed eligible and included in the review (Figure 1).

Figure 1. PRISMA diagram



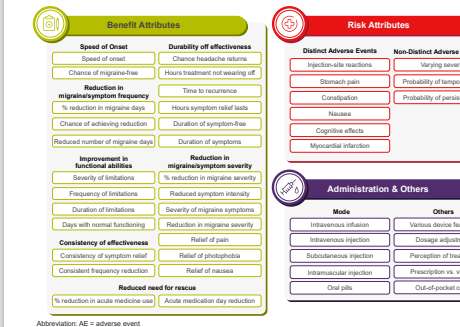
Study Characteristics

- Studies focused on preventive treatments (n=12), acute treatments (n=1) or both (n=2).
- Stated preference methods comprised discrete choice experiment (n=12), conjoint analysis (n=1), CVM (n=3), thresholding (n=1), and time trade-off (n=1).
- 17 studies derived preferences from patients; two also included clinicians (n=1) or general population (n=1).

Analytical Approaches

- The most common analysis model was mixed logit model (n=7), followed by descriptive (n=5), latent class logit model (n=3), interacted mixed logit (n=3), multinomial logit (n=2), and Hierarchical Bayes (n=2).
- Eleven studies reported preference heterogeneity by treatment experience (n=3), migraine burden (n=2), and educational attainment (n=1), suggesting the need for adequate sample size.
- The most frequent outcomes were marginal utility (n=10) and relative importance of attributes (n=10), followed by willingness-to-pay (n=5) and predicted choice/share (n=4).

Figure 2. Conceptual map of identified attributes



Methods for Attribute Development

- Five studies (all pre-2019) did not specify a method.
- Three used a single approach—literature review (n=2) or expert consultation (n=1). Ten studies used multi-method designs combining literature review with expert consultations (n=3), patient interviews or focus groups (n=3), or both (n=4).

Attribute Framing

- The average number of attributes per study was 4.9, ranging from 2–17.
- Benefit attributes included duration of relief (n=7), reduction in migraine frequency (n=8), speed of onset (n=5), reduction in migraine severity or pain (n=5), impacts on physical activities (n=5), and use of acute migraine treatments (n=1). Risk attributes included side effects categorized by reversibility (n=3) and severity (n=2), and various specific side effects such as constipation/injection site reaction (n=3). Other attributes included mode (n=5), frequency (n=5) and location (n=2) of administration, monthly cost (n=2), type of administrator (n=1), specific device features (n=1). This is summarised with Figure 2.

Choice Task Design

- Most of attributes were presented using categorical formats, including benefits (n=15), risks (n=8), and administration attributes (n=23). Among benefit attributes, percentage (n=10) and duration (n=11) formats were also common, while ratio (n=4) and frequency (n=3) were used less frequently.
- Illustrations (n=9), icon arrays (n=7), calendars (n=5), and bar charts (n=5) were the most frequently used visual aids among all attributes. However, most attributes did not use any visual aids (n=31) (Figure 3).

Attribute Ranking

- The analysis showed that benefit concepts were generally deemed most important by respondents (Figure 4). Durability of effectiveness (1.8) and consistency of effectiveness (average ranking = 2.0) were top priorities.
- For risk sub-concepts, average rankings ranged from 3.5 for gastrointestinal effects to 5.0 for injection site reactions. Mode and frequency of administration (3.0) were more influential than several risk sub-concepts, while administration setting received the lowest average rank (5.5).

Figure 3. Presentation of Attributes

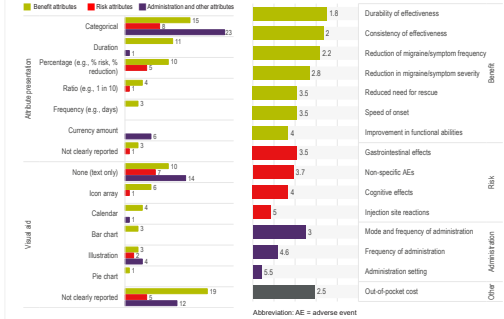
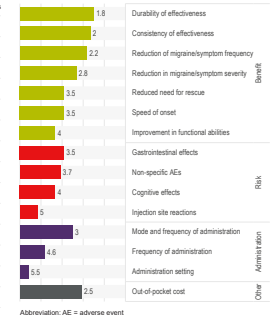


Figure 4. Average Attribute Ranking



Conclusions

- This systematic review highlights substantial variation in the design and reporting of migraine treatment preference studies. The lack of standardized methods for attribute development, framing and presentation limits comparability across studies.
- Future studies adopting transparent, best-practice methods for attribute development, framing, and analysis can strengthen the application of patient preference evidence in migraine research.

Disclosures

Dr. Seo, Dr. Clarke, and Dr. Seo are employees of PPD™ Evidera™ Patient-Centered Research, Thermo Fisher Scientific. Dr. Seo is an employee of Open Health. This poster was funded by Thermo Fisher Scientific. Editorial and graphic design support was provided by Caroline Cole and Kaveri Nayakuni of Thermo Fisher Scientific.

Please scan this QR code with your smartphone camera or app to review the published manuscript. Information obtained through digital download is for personal use only and may not be reproduced without permission from ISPOR EU or the authors of this poster.



Poster Title

Integrating Real-World Evidence into Oncology Health Technology Assessment Submissions: Recent EU Examples

Objective

- The aim of this research was to review RWE included in oncology HTA submissions in selected European countries and to identify barriers to its successful application.

<p>AUTHORS</p> <p>Ivana Sestak¹ Janneke Luijken² Fang-Yi Lin³</p> <p>Sara Angleman⁴ Alice Rouleau⁵</p>	<p>AFFILIATION</p> <p>¹Thermo Fisher Scientific, Milan, Italy ²Thermo Fisher Scientific, Ede, Netherlands ³Thermo Fisher Scientific, Waltham, MA, USA ⁴Thermo Fisher Scientific, Stockholm, Sweden ⁵Thermo Fisher Scientific, Paris, France</p>	<p>PRACTICE AREA</p> <p>PPD™ Observational Studies, Thermo Fisher Scientific</p>	<p>CLIENT</p> <p>INTERNAL</p>
---	---	--	--------------------------------------

Integrating Real-World Evidence into Oncology Health Technology Assessment Submissions: Recent EU Examples



HTA209

Ivana Sestak,¹ Janneke Luijken,² Fang-Yi Lin,³ Sara Angleman,⁴ Alice Rouleau⁵
¹Thermo Fisher Scientific, Milan, Italy; ²Thermo Fisher Scientific, Ede, Netherlands; ³Thermo Fisher Scientific, Waltham, MA, USA;
⁴Thermo Fisher Scientific, Stockholm, Sweden; ⁵Thermo Fisher Scientific, Paris, France

Background

- RWE in oncology health technology assessments (HTAs) is increasing in Europe to address limitations in traditional clinical trial data, particularly in single-arm studies, rare cancers, or accelerated approvals.
- The Joint Clinical Assessment (JCA) was implemented in the EU starting January 12, 2025, aiming to streamline health technology assessments for new cancer medicines and medical devices (as of 30 September 2025 nine applications are under review).

Objectives

- The aim of this research was to review RWE included in oncology HTA submissions in selected European countries and to identify barriers to its successful application.

Methods

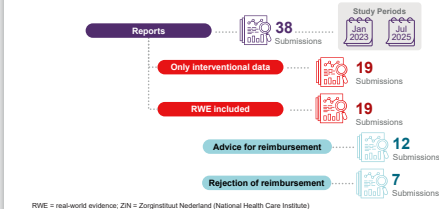
- A targeted review of oncology HTA submissions from January 2023 to July 2025 in three European countries (Netherlands, Sweden, and United Kingdom) was conducted.
- None of the included HTA submissions fall under the new JCA framework (i.e., not new cancer medicines).
- Relevant data was extracted from the submissions in the reports, including reimbursement recommendation decisions, type of RWE studies, and reasons for rejection. This data was summarized into a standardized data extraction form.
- Standard of care (SoC) data is defined here as RW data that either stem from prospective or retrospective studies and was not clearly defined in the assessment reports.
- The data are summarized with descriptive statistics, by presenting the number and proportion of advice reports in different categories.

Results

The Netherlands

- Of the 38 submissions, 19 (50.0%) included RWE, and 19 (50.0%) included only randomized-controlled trial (RCT) data.
- Twelve out of 19 submissions (63.2%) were given advice to be reimbursed (Figure 1).
- Nine submissions (47.4%) included data from retrospective studies, six (31.6%) used real-world data as external control arms, three (15.8%) used data from prospective studies, and one did not describe the type of data.
- Of the seven submissions for which reimbursement was rejected, "lack of a direct comparison" was the most common reason for rejection, followed by "issues with patients' representativeness."

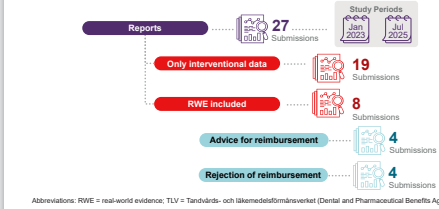
Figure 1. Zorginstituut Nederland Submissions



Sweden

- Of the 27 submissions, eight (29.6%) included RWE, and 19 (70.4%) included only RCT data.
- Fifty percent of the submissions (4/8) including RWE received advice for reimbursement (Figure 2).
- Three submissions (37.5%) included data from prospective studies, two (25.0%) included SoC data, and the remaining three included either retrospective data or data from a registry/database.
- Of the seven submissions for which reimbursement was rejected, "lack of a direct comparison" was the most common reason for rejection, followed by "uncertainty in the analysis/methodology."

Figure 2. TLV Sweden Submissions



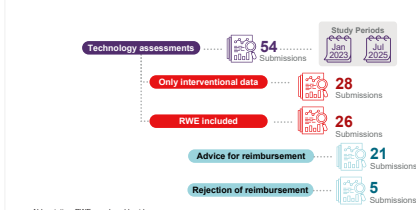
Abbreviations: RWE = real-world evidence; TLV = Tandvårds- och läkemedelsförmånsverket (Dental and Pharmaceutical Benefits Agency)

Results (cont.)

United Kingdom

- Of the 54 submissions, 26 (48.1%) included RWE, and 28 (51.9%) included only RCT data.
- Twenty-one of the 26 (80.8%) submissions that included RWE were approved for reimbursement (Figure 3).
- Nine (34.6%) submissions included SoC data, five (19.2%) used retrospective data, five included meta-analyses, and two did not describe the type of data.
- Of the five submissions for which reimbursement was rejected, "immaturity of data" and "heterogeneity of populations" were reasons stated for rejection.

Figure 3. National Institute for Health and Care Excellence Submissions

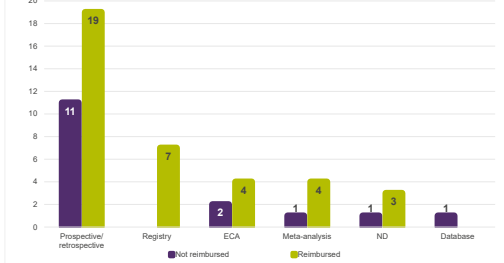


Abbreviation: RWE = real-world evidence

Summary of Submissions from All Countries

- RWE supplemented RCT data but was not used as primary evidence in any HTA submissions.
- Irrespective of reimbursement status, retrospective and prospective data (n=30) were the most frequent types of RWE included in HTA submissions, followed by registry data (n=7) (Figure 4).

Figure 4. Type of RWE by Reimbursement Status



Abbreviations: ECA = external control arm; ND = not determined; RWE = real-world evidence

Conclusions

- All submissions that included RWE were complementary to trial data rather than used as standalone evidence.
- Submissions with supplementary RWE were accepted when studies employed robust methodologies and when RW data incorporated was used as direct comparators.
- Alignment with agency expectations, data credibility, and methodological rigor are essential for successful integration.
- The UK (NICE) leads in its robust and growing use of RWE in oncology appraisals, while other countries evaluated continue to rely primarily on RCT data for oncology HTAs.

Disclosures

All authors are employees of PPD™ Observational Studies, Thermo Fisher Scientific. Funding was provided by Thermo Fisher Scientific.

Acknowledgments

Editorial and graphic design support were provided by Amara Teaboud and Kavitha Nakayima of Thermo Fisher Scientific.



ISPOR EU 2025

9–12 November 2025 | Glasgow, Scotland, UK

Poster Title

Evaluating High-cost Gene Therapies for Non-cancer Conditions: Insights From NICE's Standard Appraisal Process

Objective

- To assess the factors considered during NICE's appraisals of gene therapies for non-cancer conditions via standard STA and to determine whether the Innovative Medicines Fund (IMF) was employed to enable NICE's recommendations for high-cost, non-cancer therapies that do not meet the criteria for HST.

AUTHORS	AFFILIATION	PRACTICE AREA	CLIENT
Naomi Stapleton Alexandra Kumichel Maria Dvorkina	Thermo Fisher Scientific, London, UK	PPD™ Evidera™ Health Economics & Market Access, Thermo Fisher Scientific	INTERNAL

Evaluating High-cost Gene Therapies for Non-cancer Conditions: Insights From NICE's Standard Appraisal Process

Thermo Fisher SCIENTIFIC

HTA132

Naomi Stapleton,¹ Alexandra Kumichel,¹ Maria Dvorkina¹
¹Thermo Fisher Scientific, London, UK

Background

- Gene therapies offer significant clinical potential, yet these treatments pose substantial reimbursement challenges of high prices and uncertain long-term results.
- The National Institute for Health and Care Excellence (NICE) has typically appraised these therapies via either:
 - The highly specialised technology (HST) route which employs higher cost-effectiveness thresholds and lower evidence requirements for technologies that are innovations for an ultra-rare and debilitating condition and are likely to offer substantial additional benefit over the inadequate existing established clinical management
 - The single technology appraisal (STA) route and the Cancer Drugs Fund (CDF) for cancers not meeting HST criteria
- However, NICE has recently recommended gene therapies for non-cancer conditions via standard STA, despite list prices being >£1.6 million.

Objectives

- To assess the factors considered during NICE's appraisals of gene therapies for non-cancer conditions via standard STA and to determine whether the Innovative Medicines Fund (IMF) was employed to enable NICE's recommendations for high-cost, non-cancer therapies that do not meet the criteria for HST

Methods

- A targeted review was conducted of NICE STA guidance on gene therapies for non-cancer conditions published up to June 2025.

Results

- As of June 2025, NICE had appraised three gene therapies for non-cancer conditions through the STA route: Hemgenix for the treatment of moderately severe or severe haemophilia B and Casgevy for the treatment of transfusion-dependent beta-thalassaemia (TDT) and severe sickle cell disease (SCD).¹⁻³
- In the three appraisals, Hemgenix and Casgevy (TDT and SCD) were not recommended for routine use by NICE, but were recommended for managed access via the IMF with confidential discounts, allowing patients to access these treatments under specific conditions and monitoring.¹⁻³

Table 1. Overview of NICE STA appraisals for gene therapies for non-cancer conditions

	Hemgenix ¹	Casgevy ²	Casgevy ³
NICE STA	TA989	TA1003	TA1044
Indication	Moderately severe or severe haemophilia B	TDT	SCD
List price	£2,600,000/course	£1,651,000/course of treatment	£1,651,000/course of treatment
Initial appraisal	July 2023	N/A	March 2024
Final appraisal	July 2024	August 2024	February 2025
Recommendation	Managed access via IMF ⁴	Managed access via IMF ⁴	Managed access via IMF ⁴

Abbreviations: IMF = Innovative Medicines Fund; N/A = not applicable; NICE = National Institute for Health and Care Excellence; SCD = severe sickle cell disease; STA = single technology appraisal; TDT = transfusion-dependent beta-thalassaemia
¹Collection of additional data within 5 years through the ongoing clinical trial programme and clinical practice

Outcomes across NICE STA appraisals

- Cost-effectiveness estimates**
 - The committee noted uncertainties in cost-effectiveness estimates for all three indications, due to limited long-term data, modelling uncertainties, and incremental cost-effectiveness ratio (ICER) estimates exceeding preferred thresholds (Table 2)
 - Hemgenix:** The committee noted the treatment has the potential to be cost effective compared with congenital factor X (FIX) prophylaxis (with several scenarios falling in the acceptable range). However, the cost-effectiveness estimates were highly uncertain due to uncertainty around indirect treatment comparison (ITC) results and durability extrapolation (lack of long-term data).¹
 - Casgevy (TDT and SCD):** In both appraisals, the ICERs were considered above the preferred cost-effectiveness range (>£20,000 per quality-adjusted life-year [QALY]) in a pessimistic scenario and below the range in an optimistic scenario; uncertainty around most of the committee-preferred modelling assumptions (e.g., model structure, mortality rate, non-reference case discount rate criteria, etc.) was noted.²

Disclosures

NS, AK, and MD are employees of PPD™ Evidera™ Health Economics and Market Access, Thermo Fisher Scientific. This poster was funded by Thermo Fisher Scientific.

Acknowledgments

Editorial and graphic design support were provided by Caroline Cole and Kawahar Nakayama of Thermo Fisher Scientific.

Presented at ISPOR —The Professional Society for Health Economics and Outcomes Research—Europe | 9–12 November 2025 | Glasgow, Scotland, UK

Results (cont.)

Outcomes across NICE STA appraisals

Clinical evidence

- The NICE STA of clinical evidence indicated that, although clinical trial data for Hemgenix (haemophilia B) and Casgevy (TDT and SCD) suggested clinical effectiveness, uncertainties remained due to the lack of direct comparisons, small sample sizes, and questions about the long-term treatment effect (Table 2).¹⁻³

Equality considerations

- Equality considerations, including health inequalities related to ethnic background, as well as the innovative nature and complex technology of Casgevy, led NICE to accept higher uncertainties and consider higher cost-effectiveness estimates than usual (up to £35,000 per QALY for SCD), while no such adjustments were made for Hemgenix.¹⁻³

Table 2. Overview of outcomes assessed by NICE

	Hemgenix ¹	Casgevy (TDT) ²	Casgevy (SCD) ³
Unmet need	○ Solution to reduce the burden of prophylaxis treatment	✓ High unmet need for an effective treatment that improves outcomes for patients and QoL for patients and families/careers	✓ High unmet need for an effective, well-tolerated treatment as current treatments offer only temporary relief and do not address the underlying cause of SCD
Clinical evidence	● Single-arm trial, uncertainties regarding the ITC	● Single-arm trial, uncertainties regarding the ITC	● Single-arm trial
Long-term durability	⊖ Insufficient long-term data (up to 36 months of follow-up presented at consultation)	⊖ Insufficient long-term data (n=4259 with >16 months of follow-up)	⊖ Insufficient long-term data (up to 2 years with an average of 20.1 months of follow-up)
Cost-effectiveness estimates	● Potential for cost effectiveness, but uncertainties regarding long-term durability may result in ICERs exceeding acceptable thresholds	● Potential for cost effectiveness, but uncertainties regarding long-term durability, QoL, and outcomes in people treated with SoC may result in ICERs exceeding acceptable thresholds	● Potential for cost effectiveness, but uncertainties regarding long-term durability, QoL, and outcomes in people treated with SoC may result in ICERs exceeding acceptable thresholds
Equality considerations	○	✓ Higher uncertainty accepted due to existing health inequalities and technology's innovation and complexity	✓ Outcomes uncertainty and higher ICERs were accepted due to existing health inequalities for people with SCD

Legend: ✓ Positive effect in appraisal; ⊖ Negative effect in appraisal; ● Potentially negative effect in appraisal; ○ Not considered in appraisal
Abbreviations: ICER = incremental cost-effectiveness ratio; ITC = indirect treatment comparison; QoL = quality of life; SCD = severe sickle cell disease; SoC = standard of care; TDT = transfusion-dependent beta-thalassaemia

Managed access granted after previous rejection:

- Both Hemgenix and Casgevy (for SCD) were initially not recommended by NICE before eventually being recommended for managed access via the IMF.
 - Hemgenix:** Due to uncertainties, the treatment was not recommended for routine use; the lack of a managed access proposal during the initial submission to NICE led to no recommendation for managed access.⁴
 - Casgevy for SCD:** Due to the initial managed access proposal not addressing many of NICE's identified uncertainties and the treatment potentially not being cost-effective at the suggested price, managed access was not recommended at first.⁵

Conclusions

- High-cost gene therapies can be recommended by NICE via standard STA under managed access even after an initial negative recommendation, if there is plausible cost-effectiveness potential and new evidence could feasibly be collected to address enough uncertainties.
- The innovative nature of the technology and existing health inequalities in the target patient populations allowed higher-than-usual cost-effectiveness estimates to be considered/
- Further analysis of the appraisal process is limited due to lack of transparency in reporting cost-effectiveness estimates and discounting agreements.

References

- NICE. Safety, efficacy, and cost-effectiveness of hemgenix for treating moderately severe or severe haemophilia B. 2024. <https://www.nice.org.uk/guidance/ta989/resources/haemgenix-approval-views-for-treating-moderately-severe-or-severe-haemophilia-b-pdf40161616786789>
- NICE. TA1003. Eragapigene autotemul for treating transfusion-dependent beta-thalassaemia in people 12 years and over. 2024. <https://www.nice.org.uk/guidance/ta1003/resources/eragapigene-autotemul-for-treating-transfusion-dependent-beta-thalassaemia-in-people-12-years-and-over-pdf207362845540017>
- NICE. TA1044. Eragapigene autotemul for treating severe sickle cell disease in people 12 years and over. 2025. <https://www.nice.org.uk/guidance/ta1044/resources/eragapigene-autotemul-for-treating-severe-sickle-cell-disease-in-people-12-years-and-over-pdf207362845540017>
- NICE. TA1003. Draft guidance consultation: Eragapigene autotemul for treating moderately severe or severe haemophilia B. 2023. <https://www.nice.org.uk/guidance/ta1003/documents/ta1003-ga-consultation>
- NICE. TA1044. Draft guidance consultation: Eragapigene autotemul for treating severe sickle cell disease in people 12 years and over. 2024. <https://www.nice.org.uk/guidance/ta1044/documents/ta1044-ga-consultation>

Presented at ISPOR —The Professional Society for Health Economics and Outcomes Research—Europe | 9–12 November 2025 | Glasgow, Scotland, UK

© 2025 Thermo Fisher Scientific Inc. All rights reserved. All trademarks are the property of Thermo Fisher Scientific and its subsidiaries unless otherwise specified. This information is not intended to encourage use of these products in any manner that might violate the intellectual property rights of others.

ppd



Poster Title

Enhancing Targeted Literature Reviews (TLRs) with Artificial Intelligence (AI): A Methodological Approach for Conducting Efficient Targeted Searches

Objective

- To describe the approach used for TLRs on a variety of topics using Advancement Probability and estimate time-savings.

AUTHORS

Caroline von Wilamowitz-Moellendorff¹
Priscilla Wittkopf¹
Rajat Goel²
Sarah Ronnebaum³

AFFILIATION

¹Thermo Fisher Scientific, London, UK
²Thermo Fisher Scientific, Mumbai, India
³Thermo Fisher Scientific, Waltham, MA, USA

PRACTICE AREA

PPD™ Evidera™ Health
Economics & Market Access,
Thermo Fisher Scientific

CLIENT

INTERNAL

Enhancing Targeted Literature Reviews (TLRs) with Artificial Intelligence (AI): A Methodological Approach for Conducting Efficient Targeted Searches

Thermo Fisher
SCIENTIFIC

SA38

Caroline von Wilamowitz-Moellendorff,¹ Priscilla Wittkopf,¹ Rajat Goel,² Sarah Ronnebaum,³

¹Thermo Fisher Scientific, London, UK; ²Thermo Fisher Scientific, Mumbai, India; ³Thermo Fisher Scientific, Waltham, MA, USA

Background

- Targeted literature reviews (TLRs) are used for numerous purposes that are integral to health economics and market access. While a protocol and well-defined research questions are arguably essential components of any TLR, there is no "gold standard" TLR methodology, and therefore a flexible approach can be taken to identify the most robust and relevant information that is available in shorter timeframes than expected for systematic reviews.
- One challenge of conducting a TLR is efficiently identifying the most relevant evidence available for a given topic. Artificial intelligence (AI) tools have been employed to address challenging and time-intensive aspects of literature reviews, including search refinement, screening prediction, data extraction, risk-of-bias assessment, and synthesis.¹ For several years, the most commonly used AI application in literature reviews has been with machine learning (ML) applications, to assist in the citation screening process.² More recently, large language models (LLMs) have been utilized for literature reviews and are mainly used for search and screening steps.³
- Our team has explored various AI features within Nested Knowledge (NK) software, including searching, screening, and tagging.^{4,5} We have now tested and successfully implemented NK's Screening Model and use of an Advancement Probability score in several TLRs across different topics and therapeutic areas.

Objectives

- To describe the approach used for TLRs on a variety of topics using Advancement Probability and estimate time-savings.

Methods

- NK's screening model is a sophisticated ML algorithm designed to efficiently identify relevant studies from extensive datasets. The model requires training decisions of at least 10 included citations and 40 excluded citations to generate probability scores for each citation to advance. Each study is assigned an advancement probability score by the model, between 0 (unlikely to be relevant) and 1 (highly likely to be relevant).
- Effectiveness of the screening model is assessed using precision, recall, and accuracy scores (Table 1).

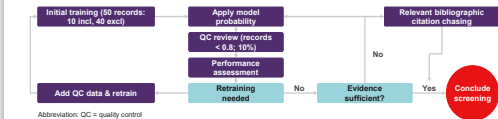
Table 1. Effectiveness definitions

Parameter	Definition
Precision	Proportion of relevant studies among those identified by the model $\frac{\text{True Positives}}{\text{True Positives} + \text{Predicted False Positives}}$
Recall	Proportion of relevant studies that the model successfully identified from the total number of relevant studies available. An ideal recall is 100%.
Accuracy	$\frac{\text{Predicted True Positives} + \text{True Negatives}}{\text{True Positives} + \text{True Negatives} + \text{False Positives} + \text{False Negatives}}$ Proportion of citations correctly identified as included or excluded
AUC	$\frac{\text{True Positives} + \text{True Negatives} + \text{False Positives} + \text{False Negatives}}{\text{True Positives} + \text{True Negatives} + \text{False Positives} + \text{False Negatives}}$ AUC is the ability to discriminate between true and false positive rates. 1 = perfect discrimination and 0.5 = no discrimination. An AUC 0.80 indicates excellent discrimination. ⁶

Abbreviation: AUC = area under the curve

- Each TLR was conducted using a protocol-driven approach with pre-specified research questions, inclusion/exclusion criteria, and search strategies. For each TLR, citations retrieved from database searches were uploaded to NK. Citations of interest were identified from a combination of keyword searches and citation-chasing from established reviews (Figure 1). After model training, researchers screened between 10 and 20 citations beginning with the citations that had the highest-advancement probability scores, checked model specifications to examine model refinement, and checked the topic coverage within the relevant articles to identify any gaps in the results.

Figure 1. Validation methodology and quality control framework



Abbreviation: QC = quality control

- Additional grey literature searches (recent congresses, clinical trial registries, and citations identified from published SLRs) were conducted to capture articles that could fill the gaps that were identified, and the model was re-run to see if new articles covering those topics were given higher advancement probability scores. This process was repeated several times until the researcher was satisfied that all research questions in the TLR had been addressed by the included literature, and that any literature "gaps" were real gaps.
- Quality control was conducted on a random sample of 15% of unscreened citations with advancement probability <math>< 0.8</math> to ensure that the most important citations were identified. All citations advanced during life/abstract screening were examined as full texts, extracted, and reported without any AI tools.

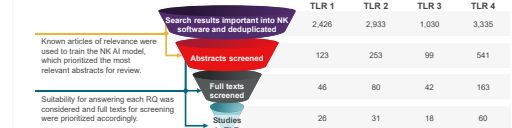
Results

- The four-step study selection approach (Figure 2) was applied to the four different TLRs.
 - Uptake, distribution of screening types, testing accuracy and chronic kidney disease (CKD) staging at diagnosis
 - Epidemiology, baseline characteristics, treatment patterns, mortality, disease progression and clinical burden of CKD due to glomerulonephritis
 - Relationship between lesion size and functional outcomes in macular degeneration
 - Clinical course, management, and patient experience of tenosynovial giant cell tumor
- Across all reviews, between 5% and 16% of abstracts required screening (Figure 3).
- Across TLRs, recall range was 0.88–0.97, accuracy 0.73–0.94, and precision 0.43–0.73, indicating that the model uses a conservative approach for determining which citations should be included, resulting in effective prioritization by the tool (Figure 4).

Results (cont.)

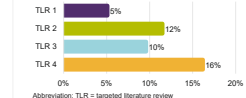
- In three of the four TLRs, some records ultimately deemed eligible had advancement probabilities below the 0.8 threshold (Figure 5), underscoring the need for regular model checks.
- Across all TLRs, we supplemented database searches with grey literature, which identified additional relevant publications, and were added to the TLR (Figure 2).
- Abstract screening time was reduced by approximately 90%, while ensuring that the most relevant studies were identified and included in the reviews; this was a consistent benefit seen across a range of topics (Table 2). Although not all articles were screened, the research questions were adequately addressed in all TLRs, demonstrating the robustness and reliability of the method.

Figure 2. Four-step study selection approach



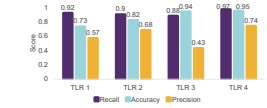
Abbreviation: NK = Nested Knowledge; TLR = targeted literature review

Figure 3. Percentage of records screened at title and abstract level



Abbreviation: TLR = targeted literature review

Figure 4. Model parameters



Abbreviation: TLR = targeted literature review

Table 2. Abstract screening time

TLR No.	Hours Required by Task	Time Saved (Hours)	% Assumptions	
	Traditional TLR ¹	AI-assisted TLR ²	Difference	
1	40.4	2.1	38.4	94.9% ³
2	36.9	4.2	32.8	88.5% ³
3	17.2	1.7	15.5	90.4% ³
4	55.6	9	46.6	83.8% ³

Abbreviation: TLR = targeted literature review

Discussion

- A cut-off score for advancement probability of 0.8 in NK's screening model was deemed to be generally appropriate for balancing precision and recall across the different topics, ensuring a high likelihood of including relevant studies while minimizing false positives. However, this threshold may need adjustment depending on the requirements of the review, such as available time for the literature review, or the acceptable level of risk for missing relevant studies.
- This approach is ideal for rapid reviews or scoping reviews where a comprehensive, but not exhaustive, search is required. Although it was shown to be suitable across the topics listed above, it is a particularly effective tool for disease areas with well-defined inclusion criteria and abundant research, such as oncology or cardiology.
- For systematic reviews, integrating a Robot Screener could be used to prioritize the most relevant abstracts for review early on. Although all abstracts still need to be double-screened by two human reviewers, prioritizing the most relevant abstracts for screening by Advancement Probability score could help to speed up the selection of highly relevant studies, and could allow reviewers to begin data extraction sooner, while articles with a lower advancement probability are screened simultaneously.
- Although this approach only saves time during abstract screening, it does result in a considerable reduction in time required during this phase of the project.
- The model's performance can continually be improved, as it automatically reassigns Advancement Probability scores to all remaining abstracts in the next after 10 abstracts have been reviewed, based on the decisions made by the human reviewer.

Conclusions

- AI-assisted screening using Advancement Probability from NK's screening model allows efficient identification of critical evidence, most suitable for targeted review methodology. The inclusion of grey literature, human validation of screening decisions, and thoughtful data interpretation and reporting are essential to support accurate evidence synthesis. Our findings suggest that AI can be a valuable tool to use for TLRs across a range of topics and therapeutic areas.

References

1. Goel R, et al. Research Synthesis Methods. 2024;15(5):531–5. Crossref: 10.1002/jrsm.1744. 2. Goel R, et al. Research Synthesis Methods. 2023;14(5):531–5. Crossref: 10.1002/jrsm.1744. 3. Goel R, et al. Research Synthesis Methods. 2023;14(5):531–5. Crossref: 10.1002/jrsm.1744. 4. Goel R, et al. Research Synthesis Methods. 2023;14(5):531–5. Crossref: 10.1002/jrsm.1744. 5. Goel R, et al. Research Synthesis Methods. 2023;14(5):531–5. Crossref: 10.1002/jrsm.1744. 6. Hosmer Jr DW, et al. Applied Logistic Regression. John Wiley & Sons, 2013.

Disclosures

All authors are employees of PPD™ Evidera™ Health Economics & Market Access, Thermo Fisher Scientific. Editorial and graphic design support was provided by Creative Circle and Kismet Creative of Thermo Fisher Scientific. This poster was funded by Thermo Fisher Scientific.

Poster Title

Cost Implications of Declining MMR Coverage in England: Modelling NHS Burden and the Value of Catch-Up Vaccination

Objective

- The primary objectives were to estimate (1) NHS and societal costs associated with reduced MMR coverage since 2019/2020; (2) the cost of delivering catch-up vaccination to unprotected children; and (3) the future economic burden if the observed decline continues.

AUTHORS

Kerry Winter
Amy Pinsent
Ruth Chapman

AFFILIATION

Thermo Fisher Scientific, London, UK

PRACTICE AREA

PPD™ Evidera™ Health Economics & Market Access, Thermo Fisher Scientific

CLIENT

INTERNAL

Cost Implications of Declining MMR Coverage in England: Modelling NHS Burden and the Value of Catch-Up Vaccination

Thermo Fisher Scientific

EE174

Kerry Winter,¹ Amy Pinsent,¹ Ruth Chapman¹
¹Thermo Fisher Scientific, London, UK

Introduction

Measles remains one of the most contagious infectious diseases, and even small declines in vaccination coverage can create significant immunity gaps.¹ Although the UK achieved measles elimination status in 2017, coverage with the measles, mumps, and rubella (MMR) vaccine has declined since the COVID-19 pandemic, particularly among 2-year-olds.¹ Current coverage is well below the 95% threshold required for herd immunity, creating immunity gaps that have already led to renewed outbreaks across the UK.² Since the start of the COVID-19 pandemic, MMR vaccination rates in England have fallen from 90.6% in 2019/2020 to 86.3% in 2023/2024 (-4.7 percentage points).² This decline increases susceptibility to outbreaks and adds avoidable healthcare costs. This study assessed the economic consequences of reduced MMR coverage for the UK National Health Service (NHS) and compared these costs with the expense of delivering catch-up vaccination to unprotected children.

Objectives

The primary objectives were to estimate (1) NHS and societal costs associated with reduced MMR coverage since 2019/2020; (2) the cost of delivering catch-up vaccination to unprotected children; and (3) the future economic burden if the observed decline continues.

Methods

A cost-of-illness model was developed to estimate the impact of declining vaccination coverage on measles cases and costs (Table 1). The model focused on the 2023 population of 2-year-old children in England (n=608,924).³ Vaccination coverage inputs were derived from NHS data,² and measles attack rates were based on UK Health Security Agency guidance (41% among unvaccinated children).⁴ Case severity distribution was assumed as 67% mild, 31% moderate, and 2% severe, informed by national data on measles cases.⁵ Direct costs included GP visits, treatment of moderate and severe measles cases, and vaccine acquisition and administration. Indirect costs captured productivity losses from caregiver absenteeism, using Office for National Statistics employment and wage data⁶ alongside Green Book assumptions⁷ on workdays missed by severity level. Analyses were conducted for both retrospective (2019–2024) and projected (2024–2030) scenarios, adopting an NHS and societal perspective. Costs are reported in 2023/24 GBP with no discounting.

Table 1. Model Inputs

Category	Parameter	Base value	Source
Population	Age 2 population (England)	608,924.00	ONS, 2023 ³
	2019/2020 MMR vaccination coverage	90.6%	NHS, 2024 ²
Vaccine coverage	2023/2024 MMR vaccination coverage	86.3%	NHS, 2024 ²
	Average annual change in MMR coverage (2019–2024)	-0.43 percentage points per year	Calculation
Epidemiology	Measles attack rate	41.00%	UKHSA, 2024 ⁴
	Proportion of mild measles cases	67.04%	GOV.UK, 2019 ⁵
	Proportion of mild/moderate cases	31.41%	GOV.UK, 2019 ⁵
	Proportion of severe measles cases	1.54%	GOV.UK, 2019 ⁵
	Cost per GP visit (mild case)	£45.00	PSSRU, 2024 ⁸
Costs	Cost per A&E visit (moderate case)	£905.00	NHS Schedule, 2023/24 ⁹
	Cost per inpatient visit (severe case)	£3,481.00	NHS Schedule, 2023/24 ⁹
	MMR vaccine acquisition cost for two doses	£15.00	BNF, 2025 ¹⁰
	MMR vaccine administration cost for two doses	£74.00	NHS Schedule, 2023/24 ⁹
Indirect costs analysis	Caregiver employment rate	75.10%	ONS, 2024 ¹¹
	Average daily wage	£143.20	ONS, 2025 ¹²
	Days of work (mild case)	2.00	Green Book (UKHSA), 2019 ⁷
	Days of work (moderate case)	4.00	Green Book (UKHSA), 2019 ⁷
	Days of work (severe case)	10.00	Green Book (UKHSA), 2019 ⁷

Abbreviations: NHS = National Health Service; ONS = Office for National Statistics; UKHSA = UK Health Security Agency.

Results

Between 2019/2020 and 2023/2024, the decline in MMR coverage resulted in an additional 10,352 unvaccinated children (Figure 1). If exposed during an outbreak in this period, this group could have generated approximately 4,244 excess measles cases. Managing these excess cases could have cost the NHS an estimated £1.16M, with an additional £1.26M in caregiver productivity losses (total £2.42M). By comparison, providing catch-up vaccination to this group would have cost £0.92M, yielding a net saving of £1.50M, making vaccination a cost-saving strategy.

If the observed decline of 0.43 percentage points per year continues, coverage is projected to fall to 86.3% by 2029/2030 (Figure 2). This would correspond to more than 83,000 unvaccinated children per birth cohort and over 34,000 measles cases per year under outbreak conditions. The cumulative six-year economic burden between 2024/2025 and 2029/2030 is projected at £79.0M, comprising £51.8M in direct NHS costs and £27.2M in caregiver productivity losses (Figure 3).

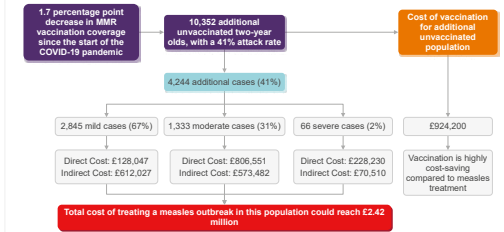
References

- UK Health Security Agency. Risk assessment for measles resurgence in the UK, 2023. <https://assets.publishing.service.gov.uk/media/64b9484820f01320cd0771a5/assessment-for-measles-resurgence-in-the-uk-2023.pdf>
- National Health Service England. Childhood vaccination coverage statistics, England, 2023/24, 2024. <https://www.nhs.uk/england/vaccination-coverage-statistics/childhood-vaccination-coverage-statistics-england-2023-24>
- Office for National Statistics. Estimates for the population of England and Wales, 2023. <https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationanddemography/datasets/estimates-for-the-population-of-england-and-wales-2023>
- Amirthalingam G, Brown K. National measles guidelines. Vol. 6.0. 2014:60.
- Public Health England. Measles cases in England, January to December 2016. https://gov.uk/government/uploads/system/uploads/attachment_data/file/481018/measles-cases-in-england-2016.pdf
- Office for National Statistics. Employment and employee types. <https://www.ons.gov.uk/peoplepopulationandcommunity/employmentandemployee/types>
- UK Health Security Agency. Measles: the green book, chapter 21, 2013.
- National Health Service England. Childhood vaccination coverage statistics, England, 2023/24, 2024. <https://www.nhs.uk/england/vaccination-coverage-statistics/childhood-vaccination-coverage-statistics-england-2023-24>
- Jones KC, et al. Unit costs of health and social care 2024 manual, 2025. <https://www.unitcosts.ac.uk/updates/170565/>
- British National Formulary. BNF Group and Pharmaceutical Press. <https://bnf.com/complete.com>
- Office for National Statistics. Average weekly earnings in Great Britain, April 2025, 2025. <https://www.ons.gov.uk/peoplepopulationandcommunity/employmentandemployee/types/earnings/bulletin/averageweeklyearningsin-great-britain/april2025>
- Chivers Moss D, et al. Vaccine 2023;14(12):4129-37.

All authors are employees of PPD™ Evidence™ Health Economics and Market Access, Thermo Fisher Scientific. This poster was funded by Thermo Fisher Scientific. Editorial and graphic design support were provided by Katherine Calves and Katherine Nakagawa of Thermo Fisher Scientific.

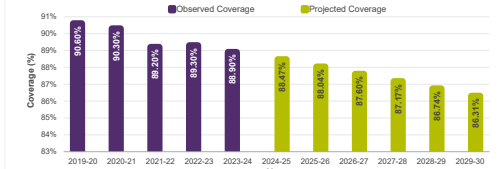
Results (cont.)

Figure 1. Model Structure for Estimating the Economic Impact of Declining MMR Coverage



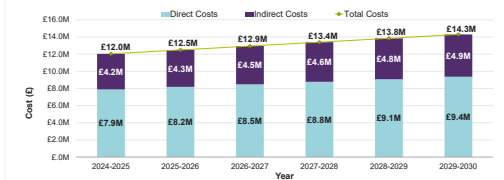
Abbreviations: MMR = measles, mumps, and rubella

Figure 2. Observed and Projected MMR Vaccination Coverage Among Two-Year-Olds in England (2019–2030)



Abbreviation: MMR = measles, mumps, and rubella

Figure 3. Projected Annual NHS and Caregiver Costs from Measles Outbreaks (2024–2030)



Abbreviation: NHS = National Health Service

Conclusions

Declining MMR coverage has already created an immunity gap that risks costly and preventable measles outbreaks. In this model, catch-up vaccination is cost-saving compared to outbreak management, as supported by UK evidence from the 2012/2013 Merseyside outbreak, where the total outbreak cost (£4.4M) far exceeded the cost of vaccinating the susceptible population (£3.18M).¹³

This pattern is also consistent with evidence that productivity losses dominate societal costs in England (74% of total when patients and carers are both included).⁸ If current trends continue, the NHS could face almost £80M in excess costs over the next six years.

These findings underscore the urgency of strengthening vaccination programmes and implementing targeted catch-up campaigns to protect public health and reduce avoidable NHS spending.

© 2025 Thermo Fisher Scientific Inc. All rights reserved. All trademarks are the property of Thermo Fisher Scientific and its subsidiaries unless otherwise indicated. This information is not intended to encourage use of these products in any manner that might infringe the intellectual property rights of others.