EMA workshop on the use of external controls for evidence generation in regulatory decision-making

3 November 2025

Summary Report

An agency of the European Union

Contents

Disclaimer

This report summarises the main discussions and insights from the workshop, aiming to support the development of a reflection paper on the use of external controls for evidence generation in regulatory decision-making. The views and reflections summarised below do not represent the official position of the European Medicines Agency (EMA) or any of its Committees or Working Parties. Nor should they be construed as reflecting a consensus amongst participants or formal outcomes of the workshop.

Executive Summary

On 3 November 2025, the EMA convened a multi-stakeholder workshop to examine the role of external controls in regulatory decision-making. Organised as a joint Methodology Working Party (MWP) and Accelerating Clinical Trials in the EU (ACT-EU) initiative, the event brought together regulators, industry representatives, academic experts, and patient organisations to discuss both the potential and the limitations of using external controls in evidence generation.

The workshop centred around four main themes:

- an overview of the current use of external controls in submissions to the EMA underlining the motivation for the drafting of a reflection paper,
- the structured pre-planning of the use of external controls including the application of the ICH E9 (R1) estimand framework and target trial emulation (TTE),
- methodological considerations for various design aspects, and
- understanding relevant parameters to identify and evaluate suitable and reliable data sources for external controls.

Workshop participants reaffirmed that randomised clinical trials (RCTs) remain the primary basis for benefit–risk assessment in the EU. External controls may provide evidence used in regulatory decision making when randomisation is unfeasible, depending on the scientific and regulatory question of interest. However, their use requires careful scrutiny due to persistent uncertainties around data quality, comparability, and residual bias. Their usability also depends on clearly defined research questions, precise estimand specifications, and rigorous assessment of data fitness for purpose—particularly in relation to representativeness, endpoint alignment, and the ability to replicate key trial features such as time zero and eligibility criteria.

The potential limitations of traditional inferential tools, including P values and confidence intervals, were noted when applied to non-randomised comparisons, reinforcing the need for transparent communication of uncertainties and structured, pre-specified decision pathways. Participants highlighted challenges in accessing high-quality comparator data, the variable reliability of real-world sources, and the importance of methodological safeguards such as TTE principles, independent oversight, and robust documentation.

Across therapeutic and population contexts—including paediatrics and rare diseases—there was broad agreement that external controls can support regulatory decisions but cannot be considered a replacement for RCT evidence. Ensuring responsible use will require continued methodological refinement, improved data availability, and strengthened communication with stakeholders, including patient communities.

The report further summarises the presentations and insights from the workshop which will support the development of the forthcoming EMA reflection paper on the use of external controls in regulatory decision making.

Welcome and Opening Remarks

The workshop was opened by **Emer Cooke**, Executive Director of the EMA and **Kit Roes**, Professor of Biostatistics at Radboud University Medical Center Nijmegen and Chair of EMA's MWP.

Emer Cooke emphasised the diversity of participants and situated the event within the broader ACT EU initiative, which aims to modernise Europe's clinical research environment through stronger, innovative trial methodologies. She stressed that early, open discussion of methodological opportunities and challenges was essential to ensuring that innovation aligns with regulatory expectations, scientific robustness, and societal needs.

She highlighted that while RCTs remain the foundation of regulatory evidence, certain contexts—such as rare diseases— might warrant alternative approaches. She concluded by encouraging active participation, framing the workshop as a step toward strengthening trust, enhancing regulatory alignment, and accelerating access to safe and effective medicines for patients.

Kit Roes stressed the strong collaboration between regulators, pharmaceutical industry, academia and patient representatives in organising the workshop. Since the recently published reflection paper on single-arm trials focuses on one data source only, addressing the use and challenges of external controls will be of added value. Reflections and guidance on this topic aligns with EMA's aspiration to strengthen evidence generation in the EU.

He emphasised the importance of understanding and formulating the research question first, and to align methodological and data source considerations to these objectives – an order that is also represented in the workshop agenda.

Peter Arlett, Head of Data Analytics and Methods Task Force (EMA) chaired the workshop together with **Kit Roes**.

Session 1: Introduction and background

Moderators: Carla Torre (Infarmed, MWP) and François Houÿez (EURORDIS)

Overview of the use of external controls in submissions to the European Medicines Agency

Johanna Lähteenvuo (Fimea, CHMP) provided an overview of when and how external controls are currently presented in EMA regulatory procedures, the limitations and expectations surrounding their use, and examples illustrating both their value and their challenges. She also highlighted methodological issues that routinely undermine the reliability of external comparisons.

The presentation emphasised that the need for an external comparison must be identified early, with plans for using external data to be made before any clinical trial results are available. The regulatory question driving the comparison must be clearly defined, and that all data collection, analyses and outcomes should be thoroughly documented and transparently reported.

Dominik Karres (EMA) provided further examples of proposals to use externa controls in paediatric studies. While randomised trials remain the preferred approach, the Paediatric Committee (PDCO) observes an increase in proposals for external controls across several therapeutic areas, particularly since the introduction of ICH E11A on extrapolation, which acknowledges that external controls can sometimes serve as a formal comparator. Their use is generally driven by the need to understand natural history and to provide a meaningful context for interpreting treatment effects in rare paediatric populations.

The presentation concluded by stressing that while external controls can play a meaningful role in paediatric development when randomisation is not possible, their use demands careful planning, early engagement with regulators and a strong understanding of the data on which the comparison will rely.

Motivation for a planned reflection paper and outline of aspects to be addressed

Elina Asikanius (Fimea, MWP) outlined MWP's motivation for developing a reflection paper on the use of external controls and how this work is being organised within the Methodology Working Party. She outlined current challenges with external control proposals, which often lack clarity, consistency and methodological robustness, and explained the need for a clearer framework to guide both applicants and assessors.

She also summarised the proposed scope of the reflection paper, including the regulatory settings in which external controls may be appropriate and the key methodological considerations that need to be addressed. The presentation concluded by stressing that the objective of the reflection paper is to improve the quality and suitability of submissions including external controls to better support regulatory decision-making.

Q&A and discussion

The Q&A session and discussion focused on the justifications for using external controls instead of RCTs, including how such justifications are reflected in regulatory documents such as the European Public Assessment Reports (EPARs). Participants explored how lack of randomisation is typically treated as an uncertainty and how and where explanations are documented. Several speakers raised the broader challenges of data suitability, the need for detailed reporting, and the limitations of current data sources—whether real-world data or historical clinical trials. The

suitability of comparator arms, the variability of data quality across sources, and the complexities of obtaining data from comparable patient populations were recurring themes.

Further questions addressed the implications of using data from different centres or time periods, concerns about representativeness, and the methodological challenges of interpreting comparative metrics when the control data has a different origin than the trial data. Participants also discussed transparency in communicating results, the constraints around using external controls in regulatory decision-making, and the need to tailor data sources to the endpoints of interest.

The session closed with remarks emphasising planning, transparency, communication, and the need for broader education—particularly for patient communities—on the methodological requirements and limitations associated with external controls.

Session 2: From concept to protocol: structured pre-planning of the use of external controls

Moderators: Murielle Mauer (EORTC) and Kit Roes (Radboud UMC, MWP Chair)

External control arms driving regulatory decision making: industry examples and learning

Robert Caroll (Bristol Myers Squibb) provided a medicine developer's perspective and outlined how external control arms are used to inform regulatory decision-making, starting with an overview of settings where they can add value. He then presented five industry case studies demonstrating different applications of external controls across therapeutic areas, highlighting their role in supporting efficacy claims, accelerating development, or addressing feasibility and ethical constraints in trial design.

The presentation concluded by emphasising the importance of early engagement with health authorities, the importance of data availability and quality as they are key to replication, adequate control of confounding and endpoint reproducibility. The relevance of approaches such as TTE and the generalisability of data across geographies were highlighted as well as the need for global alignment of guidance.

A practical framework for external control trials

Ulrich Burger (Medical University of Graz) presented a practical framework for external control trials to address potential differences between the study and external control group which could be responsible for bias.

He concluded by emphasising that external controls should not be thought of as a replacement for randomisation but as a complement. When using external controls, transparency about all relevant differences between treatment arms is key.

External controls: emulating half of the target trial

Based on causal inference as a foundation for decision-making, **Miguel Hernan** (Harvard T.H. Chan School of Public Health) introduced the concept of TTE and compared it with the ICH E9 R(1) estimand framework. His presentation focused on the application of TTE in situations where a prospective trial with external controls is planned and when external data is added at a later stage.

Using the TTE framework increases transparency in articulating the research question and how to answer it, in reporting of the results and in the evaluation of potential biases.

Discussion

The discussion centred on the applicability of methodological principles for external controls in paediatric populations, with questions raised about feasibility, ethical considerations, disease definitions, and age-specific inclusion criteria. Speakers noted that while broad methodological principles remain similar to those used in adults, paediatric subgroups introduce unique challenges linked to pharmacology, disease characterisation, and evidence thresholds. The conversation then shifted to the interpretation of P values in studies using external controls, with multiple participants questioning whether traditional statistical measures remain meaningful in the presence of uncertainty related to confounding which differ substantially from randomised settings.

Further exchanges explored how differences between databases, retrospective analyses, and variable selection processes influence statistical validity, and whether alternative approaches—such as descriptive comparisons or clearer communication of uncertainty—may sometimes be more appropriate. Participants also discussed the challenges of planning external control analyses in advance, the constraints posed by data access, and the difficulty of fully pre-specifying methods when observational data structures are not known upfront. Suggestions for potential solutions included defining decision rules rather than fixed variable lists, using independent data review committees, and improving transparency in analytic choices.

The session concluded with reflections on data access barriers, the importance of public data sources, and the need for clearer guidance on issues such as calendar time differences and acceptable data age.

Session 3: Methodological considerations for design aspects in comparisons using external controls

Moderators: Elina Asikanius (Fimea, MWP) & Patrice Verpillat (EMA)

Navigating non-randomized comparisons: from pitfalls to practice

Rima Izem (Novartis) explained why comparisons involving external control arms require careful attention to potential biases. Using a case study in rare blood disorders, she presented a structured approach to mitigate them.

The presentation concluded by highlighting that, when making a comparison not protected by randomisation, design and analytical strategies can reduce bias and that quantitative bias/sensitivity analyses can assess the impact of assumption violation. However, when and what to pre-specify in methods often still remains a challenge.

Definition of Time-Zero for Time-to-Event Data in external arm studies

Laurent Azoulay (McGill University) explained how defining time zero is central to designing external control arm studies that aim to emulate a target trial. Time zero marks the point at which follow-up begins, and individuals first become at risk for the outcome. In a randomised trial, this point is fixed by randomisation, but in external control arms, it must correspond to the moment when randomisation would have occurred in the hypothetical target trial. The presentation outlined common time-zero choices, sources of misalignment, time-related sources of bias, design and analytical approaches that respect time zero, as well as diagnostics and sensitivity analyses which were illustrated with examples.

He concluded that time-zero is a design parameter, that its correct specification is essential for causal validity and that misspecification directly leads to biased treatment effect estimates.

Quantifying and mitigating measurement bias in real-world endpoints when constructing external control arms

Benjamin Ackerman (J&J) addressed how differences in endpoint measurement between trials and real-world data can bias external control analyses, especially for time-to-event outcomes such as progression-free survival. He illustrated how to quantify bias and evaluate mitigation approaches using simulation studies and presented methods that may mitigate endpoint measurement error bias in real-world data comparators.

The presentation concluded with emphasising that robust statistical methods can adequately align real-world data endpoints towards trial standards and reduce biases for regulatory approval of external control arms.

Strategies to overcome challenges of eligibility criteria for historical data as External Control Arm source

Natalia Muehlemann (Cytel) & **Alfredo Farjat** (Bayer) presented how challenges in eligibility criteria can undermine the comparability requirements for external control studies, and why identifying both explicit and implicit challenges is essential. Steps to mitigate challenges and risk of bias were outlined and the presentation concluded with an illustrative use case of a post authorisation study.

Discussion

The panel consisting of **Angelika Geroldinger** (AGES), **Andrei Barbulescu** (EMA), **Olga Kholmanskikh Van Criekingen** (FAMHP), **Miguel Hernan** (Harvard T.H. Chan School of Public Health), and **David McConnell** (NCPE) opened with reflections on the presentations, focusing on how different methodological challenges map onto the TTE framework. Speakers distinguished between issues related to defining the estimand—such as eligibility criteria, time zero, monitoring intensity, and outcome ascertainment—and those related to confounding arising from the absence of randomisation. Additional comments highlighted the importance of pre-specification, with several panellists noting that although full pre-planning is often impossible when working with observational data, structured rules and decision frameworks can guide analytic choices before looking at the data. The discussion also covered the role of causal diagrams, the conceptual limits of adjusting for confounders, and the difficulty of aligning eligibility criteria between single-arm trials and external real-world comparators.

Perspectives from Health Technology Assessments (HTA) underscored that external controls are frequently used after marketing authorisation, but their quality varies widely and many examples are unsuitable for reliable effect estimation. Panel members emphasised the need for robust planning, transparency, and acknowledgment of residual uncertainty—especially when unmeasured confounding may strongly influence outcomes. Discussions also touched on challenges in defining follow-up time in data from existing Real-Word Data (RWD) sources, differences between real-world and clinical-trial data sources, and the limitations of constructing external controls when eligibility or endpoint definitions cannot be replicated.

The panel further explored issues around selecting external control arms, potential selection bias when there is a choice of multiple available RCTs, and questions about the feasibility of external controls in settings where treatment effects are small or patient populations are highly selected. Participants debated scenarios where external controls may be most appropriate, including cases of failed recruitment or populations poorly represented in RCTs.

The session concluded with reflections on data-quality expectations, the importance of aligning research questions with data capabilities, and the need for clearer methodological rules around data inspection and analytic decision-making when emulating target trials.

Session 4: Identification and evaluation of data sources in medicines development

Moderators: Olaf Klungel (Utrecht University) and Juan Jose Abellan (EMA)

Application of the EMA Data Quality Framework to assess the feasibility of a non-interventional study

Mònica Sabaté (VHIR) described how the EMA Data Quality Framework was translated into a practical tool to assess whether real-world data sources are suitable for use in non-interventional studies (NIS) and illustrated its application to a case study.

She concluded with general learnings and considerations and described the challenges researchers faced when using such a new tool.

Non-interventional studies: from data collection to reliable insights

Marianne Uguen (Roche) outlined the fundamentals of NIS, beginning with definitions and regulatory context, followed by the ethical and scientific standards required for generating reliable real-world evidence. She introduced the importance of data quality, common challenges in real-world data, and key mitigation strategies. A detailed case study was used to illustrate how rigorous NIS data served as a validated comparator arm in regulatory settings.

Adopting the rigorous standards of an interventional study is key to maximise the value and validity of a non-interventional study. Special emphasis was put on data quality, reliability and relevance.

Examining the Validity of External Controls Relative to Randomized Controls: The Essential Role of Data Quality

Ruthanna Davi (Medidata) presented an overview of the validity of external controls compared with randomised controls, introducing challenges of selection and confounding bias and highlighting the role of high-quality data. She outlined three oncology case studies, each illustrating how propensity-score methods and careful data preparation can align external control results with those from randomised trials.

She concluded with highlighting that selection and confounding biases can be reduced sufficiently for external controls to mimic randomised controls, that data selection and quality play a key role and that propensity score methods can substantially reduce confounding.

Discussion

The discussion focused on practical challenges and methodological considerations when using external controls, drawing on the three presentations that illustrated approaches ranging from repurposing real-world data to collecting high-quality data prospectively in non-interventional studies and using historical clinical trial data for validation. Participants asked about the limits of similarity required for matching, feasibility thresholds for propensity-score balancing, and the impact of data availability—particularly when real-world datasets lack detailed clinical outcomes such as progression-free survival. Speakers explained that historical clinical trials often provide more comparable patient populations than real-world data sources, though the ability to match depends heavily on sample size, endpoint definitions, and algorithmic validity.

Discussions also covered whether methods demonstrated in well-studied indications transfer to settings where RCTs are difficult or impossible to conduct, the challenges of applying validated

algorithms across heterogeneous clinical environments, and the role of baseline versus time-dependent confounders in intention-to-treat versus per-protocol estimands. Additional questions explored practical issues such as validating patient-reported endpoints, limitations of real-world data for measuring granular outcomes, and the iterative process of feasibility assessment, including how changes to design assumptions should be documented transparently.

The conversation closed with reflections on defining comparators appropriately—whether standard of care, placebo proxies, or untreated groups—and on the broader feasibility of applying quality-assessment frameworks to multinational and heterogeneous datasets.

Closing session and next steps

Peter Arlett and **Kit Roes** shared final reflections on the day's extensive discussions about external controls, highlighting both the methodological challenges and the opportunities they present.

Kit Roes noted a shared, nuanced view: while external controls cannot replace randomised controlled trials, they may complement them when used rigorously.

Key themes emerging during the day included the importance of clearly defining the research question and the regulatory use case, systematically assessing data quality, and applying target trial emulation principles.

Participants also emphasised persistent gaps—such as quantifying the additional uncertainty introduced by external data, limitations of P values and confidence intervals to account for such additional uncertainty in this context, and the difficulty of full pre-specification when analyses depend on observational datasets. Suggestions such as defining structured decision rules and involving independent oversight committees were mentioned as practical ways to enhance transparency and mitigate biased inference due to external data selection.

The discussions also underscored concerns around data access and data quality, noting that high-quality comparator data often comes from clinical-trial control arms that are not readily accessible, while widely available real-world data frequently lacks the necessary robustness.

Peter Arlett closed the meeting by stressing the need for continued scientific collaboration, careful methodological development, and improved data availability to ensure that external controls—when appropriately applied—can support regulatory and HTA decision-making. The workshop concluded with acknowledgments to contributors and an affirmation of the commitment to advancing evidence generation to deliver better medicines to patients.

Acronyms

ACT EU	Accelerating Clinical Trials in the European Union
AGES	Austrian Agency for Health and Food Safety
CHMP	Committee for Medicinal Products for Human Use
DQF	Data Quality Framework
ECA	External Control Arm
EMA	European Medicines Agency
EPAR	European Public Assessment Report
EORTC	European Organisation for Research and Treatment of Cancer
EURORDIS	European Organisation for Rare Diseases
FAMHP	Federal Agency for Medicines and Health Products (Belgium)
Fimea	Finnish Medicines Agency
HTA	Health Technology Assessment
ICH	International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use
MWP	Methodology Working Party
NIS	Non-Interventional Study
PDCO	Paediatric Committee
RCT	Randomised Controlled Trial
RWD	Real-World Data
RWE	Real-World Evidence
TTE	Target Trial Emulation
UMC	University Medical Center
VHIR	Vall d'Hebron Institute of Research