Externally controlled clinical trials for pediatric rare diseases: Current practices and guidance

Gabrielle Pratt Tremblay

Supervisor: Dr. Beth Potter
Co-supervisor: Dr. Christopher Gravel

Thesis submitted to the University of Ottawa in partial fulfillment of the requirements for the Master of Science degree in Epidemiology

School of Epidemiology and Public Health (SEPH)
Faculty of Medicine
University of Ottawa

Preface

Approvals

This study relied exclusively on publicly available, anonymized and/or published aggregate data and did not involve direct interaction with patients nor access to participant-level data. Therefore, it is exempt from ethics review under the Tri-Council Policy Statement (TCPS-2, Article 2.2)[1] and no ethics approval was required.

Author Contributions

The student (G.P.T) was the primary author of both manuscripts included in this thesis. Both manuscripts were co-authored by her supervisors, Dr. Beth Potter (B.K.P) and Dr. Christopher Gravel (C.G) and her thesis advisory committee member Dr. Alex Bliu (A.B). Additional authors on the first manuscript included Emma Iverson (E.I) and Dr. Jessie McGowan (J.M). E.I led the design, screening and data extraction for the parent living scoping review in collaboration with B.K.P, contributed to the interpretation of the scoping review and critically reviewed the first manuscript. J.M developed the OVID MEDLINE and Cochrane CENTRAL search strategies used in the parent living scoping review, contributed to the interpretation of the data and critically reviewed the first manuscript. B.K.P, C.G and A.B provided guidance throughout the research process and provided feedback on each chapter of this thesis.

Abstract

Objective: To describe the design and conduct of externally controlled trials in pediatric populations with genetic rare diseases (RD), including their alignment with regulatory and health technology assessment guidance.

Methods: We conducted a scoping review to identify and describe externally controlled trials in pediatric populations with genetic RDs, between 2014-2025. We also reviewed publicly available guidance on the conduct of externally controlled trials from leading regulatory and health technology assessment organizations and evaluated alignment of trials with this guidance.

Results: Among 67 articles reporting on externally controlled pediatric genetic RD trials, only eighteen applied an analytic method to address imbalances in baseline covariates between arms beyond age and baseline outcomes. Alignment with guidance was low due to poor comparability of outcome assessments, limited reporting of missingness, and under-utilization of sensitivity analyses.

Conclusion: Tailored guidance is needed to promote strategies that better mitigate against known sources of bias in externally controlled pediatric genetic RD trials.

Acknowledgments

I am so thankful to have had the opportunity to work closely with and learn from such an exceptional thesis advisory committee throughout the course of this project. To my supervisor Dr. Beth Potter, I am so incredibly grateful to have had such a caring and kind mentor like you over these last two years. I could not have gotten through this without your invaluable guidance and unwavering support. Your passion and dedication to rigorous research, teaching and mentorship have been a true inspiration and has pushed me to become a better researcher. You have taught me so much that I will always value and carry with me. To my co-supervisor Dr. Christopher Gravel, I have truly appreciated your continued support, encouragement and valuable advice throughout my master's studies. I have learned so much from you and have benefitted greatly from your knowledge and expertise. To my thesis advisory member Dr. Alex Bliu, thank you so much for your insights and the time you dedicated to this project. Your support and engagement throughout this project truly helped improve the quality of this research.

Furthermore, I would like to extend my thanks and recognitions to all the members of the RareKids-CAN screening team. This project could not have been possible without the numerous hours you all dedicated to screening and building this living inventory of trials. Also, a special thank you to my wonderful lab members, I feel so fortunate to have been a part of such a positive and uplifting group of young researchers, where I was always met with such kindness, encouragement and support.

Finally, I would like to deeply thank my family and friends for their unwavering support throughout the conduct of my graduate studies. To my parents, you are a source of stability and support that I know I can always rely on. Your unconditional love, encouragement and patience (which we all know you needed a lot of at times), has gotten me through this.

I could not have successfully completed this thesis and be taking the next steps in my development as an epidemiologist without you all.

Table of Contents

List of abbreviations

CDA: Canada's Drug Agency

CLN2: Neuronal Ceroid Lipofuscinosis type 2

EMA: European Medicines Agency

FDA: United States Food and Drug Administration

HAS: French Authority for Health

HC: Health Canada

HTA: Health Technology Assessment

ICH: International Council for Harmonization of Technical Requirements for Pharmaceutical for Human Use

INAHTA: International Network of Agencies for Health Technology Assessment

NICE: National Institute for health and Care Excellence

PRISMA-ScR: Preferred Reporting Items for Systematic reviews and Meta-Analysis extension for Scoping Reviews

RareKids-CAN: RareKids-CAN Clinical Trials and Treatment Network

RCT: Randomized controlled trial

RD: Rare disease

REDCap: Research Electronic Data Capture

ROBINS-I V2: Risk Of Bias In Non-randomized Studies of Interventions version 2

RWD: Real-world data

List of Tables

List of figures

List of supplementary appendices

1. Introduction

1.1 Rare diseases

While there is not a single agreed-upon definition of rare disease (RD) internationally,[2] in the European Union and typically in Canada, a RD is defined as any disease that affects fewer than 1 in 2,000 people.[3,4] Although individual RDs are uncommon by definition, they have been estimated to have a collective prevalence of 3.5 to 6 percent of the population globally[5] and are an important cause of morbidity and mortality, particularly in pediatric populations.[6] Many RDs are progressive, with important impacts for patients, their caregivers and health care systems. To date, over 7000 different RDs have been identified.[7] RDs are complex and can be highly heterogeneous both within and across disorders. The low prevalence and high heterogeneity of individual RDs have posed several challenges in both the understanding of RDs and the availability of effective and safe treatments. Despite these challenges, advances in biotechnology and genomics over the last decade have contributed to an increase in the development of potentially disease-modifying therapies for genetic RDs,[8] which is important given that approximately 80% of RDs have a genetic etiology.[7,9] Nonetheless, most RDs still lack an approved disease-modifying treatment.[7,10]

1.2 Challenges in conducting clinical trials for rare diseases

Clinical trials are crucial for evaluating the efficacy of new therapies and establishing their benefits to inform decisions to approve and fund them, including therapies indicated for RDs. Well-designed and conducted randomized controlled trials (RCTs), wherein participants are randomly allocated to receive the investigational treatment or an alternative treatment, placebo, or standard of care, remain the gold standard study design for evaluating the efficacy of new

treatments because they prevent confounding due to baseline differences in groups receiving the test treatment and comparator.[11] However, RCTs are often challenging to conduct for RDs, due to ethical and practical concerns.[12] Specifically, when evaluating new disease-modifying treatments that may be life-altering and there is no existing efficacious treatment, establishing an internal control group, whereby the new treatment is withheld from some participants, may be considered unethical. In addition, RCTs can be impractical for RDs due to the small number of patients eligible to participate.[13,14] In cases where an adequate internal comparison group is not feasible, investigators may conduct single-arm clinical trials, where all subjects enrolled in the trial are allocated the investigational treatment. However, by definition, single-arm trials lack a control group, do not adhere to the principles of randomization and cannot discern a comparative effect estimate. This increases the risk of confounding and other sources of bias, making the results difficult to interpret and limiting their capacity to provide causal evidence.[15]

1.3 Externally controlled trials

To partially address the limitations associated with single-arm clinical trials and attempt to characterize a treatment effect to obtain market authorization and demonstrate comparative cost effectiveness, researchers have used external sources of data to create a historical or contemporaneous comparator group.[14,16,17] While different authors and organizations use different terminology to describe the design of clinical studies, the United States Food and Drug Administration (FDA) defines externally controlled trials as trials where "outcomes in participants receiving the test treatment according to a protocol are compared to outcomes in a group of people external to the trial who had not received the same treatment".[18,p.1] Typically when a trial is externally controlled, there is no internal control group and the external control

group is the only control arm, however, external control data may be incorporated to supplement the within-trial control data. Consent to the trial with respect to the intervention under investigation is not gathered for individuals in the external control group because they were not offered the opportunity to participate in the trial or the possibility of receiving the treatment under investigation. Frequently, external control data is drawn from historical sources prior to the availability of the treatment being evaluated but external control data can also be concurrent. The decision to include an external control group in a trial can be made a priori or post-hoc, although a priori inclusion is strongly recommended and is required by most regulators.[18]

Potential data sources for external controls include control arms from other clinical trials, observational studies, such as natural history studies, or other real-world data (RWD) sources, such as patient registries, electronic health records or claims data.[19-21] RWD is defined by Health Canada (HC) and the FDA as "data relating to patient health status and/or the delivery of health care routinely collected from a variety of sources".[18,22] Typically individual patient data from the external control group are incorporated into the analysis for the clinical trial but aggregate data have also been used.[23]

Over the last decade, there has been an increase in the number of submissions to regulatory organizations, such as the FDA, the European Medicines Agency (EMA), and HC, and to health technology assessment (HTA) organizations such as Canada's Drug Agency (CDA), that incorporate results from externally controlled trials in their applications for marketing approval or reimbursement, respectively, for new treatments.[24-27] Based on guidance from these regulators and HTA organizations, and/or on publications that have summarized their reviews/feedback on submissions that included data from externally controlled trials, these organizations have recognized the potential value of this approach to evaluating treatments in the

case of RDs and where there is no alternative option for generating a control arm, but have also identified important challenges and sources of bias.[13,18,20,22,28]

1.4 Challenges in conducting externally controlled trials

Despite their promise, treatments evaluated using an external comparator often face greater challenges in receiving regulatory approval and positive reimbursement recommendations due to uncertainty about the evidence generated from these trials, with reviewers raising concerns about data quality and potential biases including unmeasured confounding, selection, and information biases, often related to the comparability of the external data sources with the associated single arm clinical trial.[6,18,20,29] Most commonly, reviewers are concerned with differences in patient population characteristics, such as the alignment of patient eligibility criteria, including age, sex and disease severity; and/or differences in time-related factors, comparability of the measurement and evaluation of trial endpoints, or missing data. To mitigate these sources of bias, various statistical techniques, such as propensity score matching, inverse probability of treatment weighting and Bayesian methods have been used.[16,30,31] However, some of these methods may be difficult to apply in RD settings due to small sample sizes and can introduce additional bias if their assumptions are not met.[30]

An illustrative example in an RD context is an externally controlled trial that was conducted to evaluate the drug, cerliponase alfa (Brineura), an enzyme replacement therapy developed to treat neuronal ceroid lipofuscinosis type 2 (CLN2), a form of Batten disease.[32] Data from the externally controlled trial was part of the submission to regulators and given its eventual market approval, some regulatory and reimbursement review reports are publicly available. Among the limitations of this externally controlled trial discussed in reports from

reviewers at the FDA and the Canadian Agency for Drug and Technologies in Health (now Canada's Drug Agency), were concerns over the comparability and convergent validity of endpoint measurements, calendar time differences, irregularity of assessment times between data sources, comparability and availability of information on baseline characteristics, and methods used for matching.[33,34] These concerns are not unique to this example. The STR1VE trial was an externally controlled trial conducted to evaluate Onasemnogene abeparvovec (Zolgensma), a gene therapy developed to treat spinal muscular atrophy type 1.[35] Data from this externally controlled trial were considered in the submission for approval and reimbursement of this treatment by the FDA and CDA. Similar to the example of Brineura, reviewers expressed concerns over the comparability of baseline characteristics across internal and external arms, limited use of analytic methods to account for these differences, calendar time differences and lack of blinding.[36,37] Despite these recurrent concerns, there remains limited publicly available regulatory or HTA guidance for clinical trialists that is specific to the conduct of externally controlled trials, particularly in the context of smaller sample sizes and other characteristics commonly encountered in pediatric genetic RD trials (e.g., slowly progressive severe diseases diagnosed in early childhood, high clinical heterogeneity). Furthermore, the limited guidance available has not been systematically summarized across agencies and organizations.

1.5 Summary of rationale

To generate meaningful evidence and accelerate advances in the management of RDs, clinical trials are essential. Externally controlled trials may be particularly well-suited to evaluating and enhancing accessibility of disease-modifying treatments for pediatric genetic RDs but are associated with important limitations that impede the robustness of the evidence they produce

and have led to challenging policy decisions about licensing and reimbursement. As this approach becomes more common with the increase in disease-modifying therapies for RDs, there is a need to understand current practices in externally controlled trial design and conduct for pediatric genetic RDs and to critically evaluate the alignment of these practices with existing regulatory and HTA guidance. This is an essential step toward recommendations for improved evidence generation, tailored guidance and ultimately better access to promising therapies for patients.

1.6 Thesis scope

This thesis focuses specifically on externally controlled trials among pediatric populations with genetic RDs. Children are disproportionally impacted by RDs, with approximately 75% of RDs manifesting in childhood.[7] Furthermore, approximately 80% of RDs have a genetic origin.[7] As noted, gene therapies and other disease-modifying interventions are emerging rapidly and have the potential to address the high unmet needs of children with genetic RDs, and RCTs are frequently considered infeasible for evaluating these therapies.

1.7 Research objectives

We aimed to describe and critically evaluate pediatric genetic RD clinical trials that have used external controls, informing recommendations for trialists and for developers of registries and natural history studies that may be sources of RWD for these trials. Ultimately, we wish to contribute to improving the evidence generated from these trials. Our specific objectives were to:

1. Identify and describe the study characteristics and methods used in pediatric genetic RD clinical trials that rely on external data as a source of controls;

2. Summarize existing, publicly available guidance and criteria published by leading regulatory agencies and HTA organizations, directed to sponsors/trialists or evaluators/reviewers of externally controlled RD trials; and
3. Critically evaluate the quality and use of external controls in the trials from objective (i), using the guidance and criteria identified in objective (ii).

1.8 Structure of thesis

This thesis is presented by manuscript, in accordance with the guidelines provided by the School of Epidemiology and Public Health at the University of Ottawa. A preface is included before each manuscript to indicate the thesis objectives it aims to address, the author contributions and journal submission information. The order is as follows:

1. Chapter 1: Provides an overview of the existing literature on the topic of externally controlled trials, rationale and objectives for this thesis.
2. Chapter 2: The first manuscript, entitled *Externally controlled trials in pediatric rare genetic disease: A scoping review*, describes the current practices and methods of externally controlled trials in pediatric genetic RDs, with a more in-depth exploration of study reports that analyzed individual patient data from the external control arm(s) and attempted to statistically address differences in baseline characteristics between the internal and external arms. This manuscript was authored by Gabrielle Pratt Tremblay, Emma Iverson, Dr. Jessie McGowan, Dr. Alex Bliu, Dr. Christopher Gravel and Dr. Beth Potter and has been prepared for submission to the Journal of the American Medical Association (JAMA) Network Open.

3. Chapter 3: The second manuscript, entitled *Guidance for externally controlled trials: A review and application to pediatric rare disease*, summarizes current guidance on externally controlled trials published by leading regulatory and HTA agencies, and assesses the application of this guidance in externally controlled trials conducted in pediatric populations with rare diseases. This manuscript was authored by Gabrielle Pratt Tremblay, Dr. Alex Bliu, Dr. Christopher Gravel and Dr. Beth Potter and has been prepared for submission to Nature Communications.
4. Chapter 4: Presents an integrated discussion, including the interpretation of the study findings across the two manuscripts and their implications for the future development of guidance for externally controlled trials.

1.9 References for Chapter 1

1. Canadian Institutes of Health Research, Natural Sciences and Engineering Research Council of Canada, Social Sciences and Humanities Research Council of Canada. Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans. 2022. Accessed January 20, 2024. https://ethics.gc.ca/eng/documents/tcps2-2022-en.pdf

2. Richter T, Nestler-Parr S, Babela R, et al. Rare disease terminology and definitions - A systematic global review: Report of the ISPOR rare disease special interest group. *Value Health*. 2015;18(6):906-914. doi:10.1016/j.jval.2015.05.008

3. European Commission. Rare diseases: EU research on rare diseases. Accessed April 1, 2025. https://research-and-innovation.ec.europa.eu/research-area/health/rare-diseases_en

4. Canadian Organization for Rare Diseases. About CORD: Key facts. Accessed February 20, 2025. https://www.raredisorders.ca/about/about-cord

5. Nguengang Wakap S, Lambert DM, Olry A, et al. Estimating cumulative point prevalence of rare diseases: Analysis of the Orphanet database. *Eur J Hum Genet*. 2020;28(2):165-173. doi:10.1038/s41431-019-0508-0

6. Rath A, Salamon V, Peixoto S, et al. A systematic literature review of evidence-based clinical practice for rare diseases: What are the perceived and real barriers for improving the evidence and how can they be overcome? *Trials*. 2017;18(1):556. doi:10.1186/s13063-017-2287-7

7. The landscape for rare diseases in 2024. *Lancet Glob Health*. 2024;12(3):e341. doi:10.1016/S2214-109X(24)00056-1

8. Baek HJ, Nam Y, Rim YA, Ju JH. Advances in cell and gene therapy for rare disease treatment. *Int J Stem Cells*. 2025;18(4):324-337. doi:10.15283/ijsc25028

9. Dawkins HJS, Draghia-Akli R, Lasko P, et al. Progress in rare diseases research 2010–2016: An IRDiRC perspective. *J Clin Transl Sci*. 2018;11(1):11-20. doi:10.1111/cts.12501

10. Fermaglich LJ, Miller KL. A comprehensive study of the rare diseases and conditions targeted by orphan drug designations and approvals over the forty years of the Orphan Drug Act. *Orphanet J Rare Dis*. 2023;18(1):163. doi:10.1186/s13023-023-02790-7

11. Schulz KF, Grimes DA. Blinding in randomised trials: Hiding who got what. *Lancet*. 2002;359(9307):696-700. doi:10.1016/S0140-6736(02)07816-9

12. Augustine EF, Adams HR, Mink JW. Clinical trials in rare disease: Challenges and opportunities. *J Child Neurol*. 2013;28(9):1142-50. doi:10.1177/0883073813495959

13. Khachatryan A, Read SH, Madison T. External control arms for rare diseases: Building a body of supporting evidence. *J Pharm Dyn*. 2023;50(6):501-506. doi:10.1007/s10928-023-09858-8

14. Jahanshahi M, Gregg K, Davis G, et al. The use of external controls in FDA regulatory decision making. *Ther Innov Regul Sci*. Sep 2021;55(5):1019-1035. doi:10.1007/s43441-021-00302-y

15. Wang M, Ma H, Shi Y, Ni H, Qin C, Ji C. Single-arm clinical trials: Design, ethics, principles. *BMJ Support Palliat Care*. 2025;15(1):46. doi:10.1136/spcare-2024-004984

16. Liu J, Barrett JS, Leonardi ET, et al. Natural history and real-world data in rare diseases: Applications, limitations, and future perspectives. *J Clin Pharmacol*. 2022;62(Suppl 2):S38-s55. doi:10.1002/jcph.2134

17. European Medicines Agency. *Committee for Medicinal Products for Human Use (CHMP) reflection paper on establishing efficacy based on single-arm trials submitted as pivotal evidence in a marketing authorisation application: Considerations on evidence from single-arm trials*. 2024. Accessed November 2, 2024. https://www.ema.europa.eu/en/documents/scientific-guideline/reflection-paper-establishing-efficacy-based-single-arm-trials-submitted-pivotal-evidence-marketing-authorisation-application_en.pdf

18. United States Food and Drug Administration. *Considerations for the design and conduct of externally controlled trials for drug and biological products guidance for industry: Draft guidance*. 2023. Accessed October 24, 2024. https://www.fda.gov/vaccines-blood-biologics/guidance-compliance-regulatory-information-biologics/biologics-guidances

19. Wu J, Wang C, Toh S, Pisa FE, Bauer L. Use of real-world evidence in regulatory decisions for rare diseases in the United States - Current status and future directions. *Pharmacoepidemiol Drug Saf*. 2020;29(10):1213-1218. doi:10.1002/pds.4962

20. Sola-Morales O, Curtis LH, Heidt J, et al. Effectively leveraging RWD for external controls: A systematic literature review of regulatory and HTA decisions. *Clin Pharmacol Ther*. 2023;114:325-355. doi:10.1002/cpt.2914

21. Seeger JD, Davis KJ, Iannacone MR, et al. Methods for external control groups for single arm trials or long-term uncontrolled extensions to randomized clinical trials. *Pharmacoepidemiol Drug Saf*. 2020;29(11):1382-1392. doi:10.1002/pds.5141

22. Health Canada. *Elements of real world data/evidence quality throughout the prescription drug product life cycle*. 2019. Accessed June 12, 2025. https://www.canada.ca/en/services/health/publications/drugs-health-products/real-world-data-evidence-drug-lifecycle-report.html#a1

23. Lambert J, Lengliné E, Porcher R, Thiébaut R, Zohar S, Chevret S. Enriching single-arm clinical trials with external controls: Possibilities and pitfalls. *Blood Adv*. 2023;7(19):5680-5690. doi:10.1182/bloodadvances.2022009167

24. Mack C, Christian J, Brinkley E, Warren EJ, Hall M, Dreyer N. When context is hard to come by: External comparators and how to use them. *Ther Innov Regul Sci*. 2020;54(4):932-938. doi:10.1007/s43441-019-00108-z

25. Thorlund K, Dron L, Park JJH, Mills EJ. Synthetic and external controls in clinical trials - A primer for researchers. *Clin Epidemiol*. 2020;12:457-467. doi:10.2147/clep.S242097

26. Wang X, Dormont F, Lorenzato C, Latouche A, Hernandez R, Rouzier R. Current perspectives for external control arms in oncology clinical trials: Analysis of EMA approvals 2016–2021. *J Cancer Policy*. 2023;35:100403. doi:10.1016/j.jcpo.2023.100403

27. Patel D, Grimson F, Mihaylova E, et al. Use of external comparators for health technology assessment submissions based on single-arm trials. *Value Health*. 2021;24(8):1118-1125. doi:10.1016/j.jval.2021.01.015

28. Curtis LH, Sola-Morales O, Heidt J, et al. Regulatory and HTA considerations for development of real-world data derived external controls. *Clin Pharmacol Ther*. 2023;114(2):303-315. doi:10.1002/cpt.2913

29. Velummailum RR, McKibbon C, Brenner DR, Stringer EA, Ekstrom L, Dron L. Data challenges for externally controlled trials: Viewpoint. *J Med Internet Res*. 2023;25:e43484. doi:10.2196/43484

30. Hogervorst MA, Soman KV, Gardarsdottir H, Goettsch WG, Bloem LT. Analytical methods for comparing uncontrolled trials with external controls from real-world data: A systematic literature review and comparison with European regulatory and health technology assessment practice. *Value Health*. 2025;28(1):161-174. doi:10.1016/j.jval.2024.08.002

31. Warden MN, Heltshe SL, Simon N, Mooney SJ, Mayer-Hamblett N, Magaret AS. Statistical approaches for the integration of external controls in a cystic fibrosis clinical trial: A simulation and an application. *Am J Epidemiol*. 2024;193(12):1796-1804. doi:10.1093/aje/kwae148

32. Schulz A, Ajayi T, Specchio N, et al. Study of intraventricular cerliponase alfa for CLN2 disease. *N Engl J Med*. 2018;378(20):1898-1907. doi:10.1056/nejmoa1712649

33. United States Food and Drug Administration. *Statistical review: NDA 761052, Brineura (cerliponase alfa)*. 2017. Accessed April 30, 2025. https://www.accessdata.fda.gov/drugsatfda_docs/nda/2017/761052Orig1s000StatR.pdf

34. Canadian Agency for Drugs and Technologies in Health. *Common drug review: Clinical review report – Cerliponase Alfa (Brineura)*. 2019. Accessed April 30, 2025. https://www.cda-amc.ca/sites/default/files/cdr/clinical/sr0574-brineura-clinical-review-report.pdf

35. Day JW, Finkel RS, Chiriboga CA, et al. Onasemnogene abeparvovec gene therapy for symptomatic infantile-onset spinal muscular atrophy in patients with two copies of SMN2 (STR1VE): An open-label, single-arm, multicentre, phase 3 trial. *Lancet Neurol*. 2021;20(4):284-293. doi:10.1016/S1474-4422(21)00001-6

36. Canadian Agency for Drugs and Technologies in Health. *Common drug review: Clinical review report – Onasemnogene Abeparvovec (Zolgensma)*. 2021. Accessed February 2, 2026.

https://www.cda-amc.ca/sites/default/files/DRR/2021/sg0649-zolgensma-clinical-review-report.pdf

37. United States Food and Drug Administration. *Statistical review: BLA 125964, Zolgensma (Onasemnogene Abeparvovec)*. 2019. Accessed February 2, 2026.

Chapter 2: Preface

Objective: To identify and describe the design and methodological characteristics of externally controlled trials in pediatric genetic RD.

Author contributions: Gabrielle Pratt Tremblay led the conception and design of this study, the screening, extraction and analysis for the scoping review, and interpretation of the data, and drafted the manuscript. She also participated in the screening and extraction of the RareKids-CAN living scoping review used as a data source for this study. Emma Iverson led the design, screening and extraction for the RareKids-CAN living scoping review, contributed to the screening and interpretation for this scoping review on externally controlled trials and critically reviewed the manuscript. Dr. Jessie McGowan participated in the acquisition and interpretations of the data and critically reviewed the manuscript. Dr. Alex Bliu and Dr. Christopher Gravel contributed to the conception, design, and supervision of the study, interpretation of the data, and critically reviewed the manuscript. Dr. Beth Potter contributed to the conception, design, and supervision of the study, screening and extraction of the eligible articles, analysis and interpretation of the data, and critically reviewed the manuscript.

Publication: This manuscript is prepared for submission to the Journal of the American Medical Association (JAMA) Network Open. Additional details in the methods and results have been added in this version to improve readability for the examiners of this thesis.

2. Manuscript 1 - Externally controlled trials in pediatric rare genetic disease: A scoping review

Gabrielle Pratt Tremblay[1], Emma Iverson[1], Jessie McGowan[1], Alex Bliu[2], Christopher Gravel[1,3,4], Beth K. Potter[1]

Affiliations:

[1]School of Epidemiology and Public Health, University of Ottawa, ON, Canada,

[2]Office of Biostatistics, Biologic and Radiopharmaceutical Drugs Directorate, Health Canada, Ottawa, ON, Canada

[3]Department of Mathematics and Statistics, University of Ottawa, Ottawa, ON, Canada

[4]Data Literacy Research Institute, University of Ottawa, ON, Canada

Funding

This study is funded by the RareKids-CAN: Pediatric Rare Disease Clinical Trials and Treatment Network. RareKids-CAN is funded by the Canadian Institutes of Health Research as part of the Government of Canada's National Strategy of Drugs for Rare Diseases (Grant #192262).

2.1 Key points

Question: How are externally controlled trials in populations with pediatric rare genetic diseases designed, conducted and reported?

Findings: Sixty-seven articles reporting on externally controlled trials were identified through a scoping review. Fifty-one articles (76%) used individual-level data to create an external control arm, however, only 18 (27%) attempted to address differences in baseline characteristics across arms. Comparability of outcome assessment time points, data quality, reporting of missingness in the external control arm and use of sensitivity analyses were generally poor.

Meaning: The findings of this review point to the need for guidance specific to externally controlled trials in genetic rare disease.

2.2 Abstract

Importance: Given that randomized controlled trials are often unfeasible in pediatric genetic rare disease (RD), externally controlled trials represent a promising alternative. However, treatments evaluated using external controls often face challenges in obtaining approval and/or reimbursement. To support the conduct of future externally controlled trials and increase access to effective disease-modifying therapies in pediatric RD, more information is needed on current practices.

Objective: To identify and describe the methodological characteristics and reporting practices of externally controlled trials in pediatric genetic RD.

Evidence Review: This scoping review followed JBI methodology and is reported according to the Preferred Reporting Items for Systematic reviews and Meta-Analysis extension for Scoping Reviews guideline. We identified externally controlled trials published between January 1, 2014, and February 26, 2025, from a previously conducted living scoping review of pediatric RD trials and supplementary searches. Eligible trials were conducted in children with genetic RDs and used external data to make a comparison to at least one efficacy or safety outcome.

Findings: Of 67 articles reporting externally controlled trials conducted across 6 continents for 37 different genetic RDs, 16 described a comparison to aggregate external data and 51 made a comparison to individual-level external data. Among these 51, 18 articles (27%) reported on trials using an analytic method to address imbalances in baseline covariates between arms, beyond age and baseline outcome measurement. The most common analytic methods were matching (n=12) and regression analysis (n=8). Of the 18 articles, most evaluated single-arm trials, pre-planned their external comparison and utilized observational cohort data as the external control. Ten of the 18 articles (56%) reported an external comparison for at least one

primary outcome. Comparability of outcome assessment time points, the quality of the data and reporting of missingness in the external control arm and the use of sensitivity analyses were generally poor.

Conclusions and Relevance: A minority of externally controlled trials in pediatric populations with a genetic RD addressed baseline covariate imbalances. The limitations we identified in methodological and reporting practices suggest a need for guidance for externally controlled trials in genetic RDs to improve the quality of evidence generated from these trials.

2.3 Introduction

In Canada and the European Union, a rare disease (RD) is commonly defined as any disease that affects fewer than 1 in 2,000 people,[1,2] and in the United States as a disease that affects fewer than 200,000 people.[3] It is estimated that there are over 7,000 different RDs, 80% of which have a genetic origin.[4] They have been estimated to affect up to 6% of the population globally and are an important cause of morbidity and mortality, particularly in pediatric populations.[5,6] Despite their collective prevalence, only approximately 5% of RDs have specific treatments available, resulting in a very high unmet need.[4]

Advances in biotechnology and genomics over the last decade have contributed to an increase in the development of potentially disease-modifying therapies for genetic RDs.[7] While randomized controlled trials (RCTs) are considered the optimal design to establish the efficacy of new and emerging treatments, they are often not feasible or ethical to conduct in populations with RDs.[8] This is due to the small sample sizes, a high clinical heterogeneity within diseases, and the lack of an alternate treatment option.[9,10] Specifically, in RD populations with no existing efficacious treatment, establishing an internal control group, whereby a potentially disease-modifying treatment is withheld from some participants, may be considered unethical.

To design studies that enable a comparator despite the infeasibility of RCTs, an increasing number of RD trialists have been conducting externally controlled trials, whereby external sources of data are used to create a comparison arm from which a treatment effect can be determined. Externally controlled trials are defined by the United States Food and Drug Administration (FDA) as trials in which "outcomes in participants receiving the test treatment according to a protocol are compared to outcomes in a group of people external to the trial who had not received the same treatment".[11,p.1] Potential sources of external control data include

previously conducted clinical trials, observational longitudinal studies, or other real-world data sources.

Many regulatory and health technology assessment organizations have acknowledged the value of externally controlled trials in circumstances in which an internal comparison group is not feasible and/or where there is a high unmet need, such as in RDs, however, they have also raised important concerns about the conduct of externally controlled trials.[11-16] These concerns, often raised at the revision stage of submissions, include the potentially poor comparability of population characteristics, prognostic factors, outcome measurement and follow-up across arms; temporal and geographic differences in diagnostic rates and health care when external control data are historical rather than concurrent and/or in a different region; and missing data in the external control arm(s). These concerns have led to challenging policy decisions about the licensing and reimbursement of certain treatments. Some organizations have released preliminary guidance on the inclusion of external data in submissions, however, current guidance is broad and not specific to the area of pediatric RDs.[11,16-20]

To increase understanding of current practices in externally controlled trials in pediatric RDs and support the future generation of more specific guidance for trialists and real-world data producers, a review of the current literature is needed. To date, reviews conducted on externally controlled trials have been broad across disease areas,[21] specific to trials submitted for regulatory or health technology assessment review,[13,14,22,23] or specific to hematological cancers or immune-related inflammatory diseases.[24,25] To our knowledge, no study has explored the design, study characteristics, and methodology of externally controlled trials specifically conducted in pediatric populations with a genetic RD, an area with a very high unmet need that faces unique challenges and concerns that externally controlled trials aim to address. This scoping review

aimed to identify and describe the characteristics of externally controlled trials in this area and examine the statistical methods used to address differences in baseline characteristics between internal and external arms using individual-level external control data.

2.4 Methods

A scoping review was chosen, given the exploratory nature of the objectives and our desire to describe and map study methods rather than to synthesize study findings. We followed the JBI methodology,[26] and report the review in accordance with the Preferred Reporting Items for Systematic reviews and Meta-Analysis extension for Scoping Reviews (PRISMA-ScR) guidelines.[27] The protocol for this scoping review is publicly available and was registered on Open Science Framework prior to screening.[28]

2.4.1 *Information sources and search strategy*

An existing inventory of articles from a living scoping review conducted by the RareKids-CAN Pediatric Clinical Trials and Treatment Network (hereafter referred to as the "RareKids-CAN inventory") was the main source of records for this scoping review.[29,30] The RareKids-CAN inventory includes pediatric RD clinical trial protocols and reports published in peer-reviewed journals from January 1, 2014, to February 26th, 2025, and was developed using JBI scoping review methodology.[26] Citations for the RareKids-CAN inventory were identified through systematic searches conducted in OVID MEDLINE and the Cochrane CENTRAL on February 26th, 2025 with the guidance of an experienced librarian (JM). The full search strategy can be found in **eMethods 1**. Following full-text screening for the present review, we identified additional articles through citation searching of the reference lists of included articles.

2.4.2 *Eligibility criteria*

Articles were eligible for inclusion in the RareKids-CAN inventory if they: 1) included children (<18 years of age) with a RD; 2) were a main trial report, protocol, feasibility study or pilot study; 3) aimed to evaluate an intervention designed to improve health-related outcomes; 4) included primary research; and 5) were published in English between January 1st, 2014 and February 26th, 2025.

Articles from the RareKids-CAN inventory were eligible for inclusion in this scoping review if they: 1) met the FDA definition of an externally controlled trial;[11] 2) ≥ 50% of study participants were diagnosed with an RD of genetic and non-oncologic origin, as indicated in the Orphanet Rare Disease Database and/or the National Institutes of Health Genetic and Rare Diseases list;[31,32] and 3) the study reported comparisons of internal trial data to external data for one or more safety or efficacy outcomes.

2.4.3 *Selection of evidence*

To create the RareKids-CAN inventory, citations retrieved by the search were uploaded to Covidence software.[33] Two trained reviewers (among a team of reviewers, including lead author G.P.T) independently screened citations in two stages (title/abstract and full-text) against the predefined RareKids-CAN eligibility criteria (**eFigure 1**). The two original screeners resolved conflicts through discussion, consulting a third reviewer when needed. For the present review, we imported all articles included in the RareKids-CAN inventory into a separate Covidence project. Two independent reviewers (including G.P.T for all articles) screened all articles at the full-text level using the eligibility criteria described above. The two original screeners resolved conflicts through discussion, involving a third reviewer (B.K.P) when needed.

2.4.4 *Data extraction and data items*

Following full-text screening, we extracted data from all articles meeting the inclusion criteria. We used a custom data extraction form using the Research Electronic Data Capture (REDCap) software,[34] developed by E.I and B.K.P. This form captured general study characteristics, including trial registration, study design, study site countries, funding, trial phase, number of internal arms, allocation method, sample size, RD, and sex, gender, and age. For articles registered on the ClinicalTrials.gov trial register and where a national clinical trial number (commonly referred to as an NCT) could be identified, selected variables were extracted directly from Clinicaltrials.gov, using the Aggregate Analysis of ClinicalTrials.gov database.[35] Two team members extracted data independently (G.P.T and E.I extracted data for all externally controlled trials that met eligibility criteria for the present review), resolving discrepancies through discussion and involving a third reviewer (B.K.P) when needed.

Following initial extraction, we made a post-hoc decision to prioritize a subset of articles for more detailed extraction due to heterogeneity in the granularity of external data, the use of statistical adjustment methods and reporting practices. Articles that satisfied all the following criteria were prioritized for detailed extraction: 1) the authors derived external control estimates from individual-level data; and 2) used an analytic method beyond adjustment for age and baseline outcome scores to address baseline imbalances between internal and external comparison groups (hereafter referred to as the “prioritized subset”). We included matching as a method for addressing baseline covariate imbalance but a challenge was that authors used the term "matching" to refer to different approaches, ranging from propensity score matching on a comprehensive set of covariates, to matching on a small number of individual variables, to group matching using specific variables and methods, to "matching" that referred to applying the same

or similar inclusion criteria in the trial and external control. Therefore, articles that used matching were included in the subset for detailed extraction only if they matched on specific variables beyond age and the baseline value of the outcome and described methods that went beyond application of the same or similar inclusion criteria.

We used a second extraction form, developed by G.P.T and B.K.P using REDCap, to capture information on the characteristics of external data sources (data type, number of arms, sample size and age of participants), comparability between internal and external arms (inclusion criteria, covariates used for adjustment, methods to address covariate imbalance and data collection periods), details regarding outcomes compared in the internal versus external arms (outcome type, measurement timing and frequency, blinding, handling of missing data, statistical methods and sensitivity analyses) and reporting characteristics (protocol availability, losses to follow-up, and adverse events), for all articles included in the prioritized subset. For each article, data were extracted from the study report, protocol, statistical analysis plan and trial registration, when available. Detailed extraction for the subset of articles that met all criteria was conducted by one team member (G.P.T) and verified by a second team member (B.K.P).

We piloted both extraction forms (**eMethods 2 and 3**) on 3 randomly selected eligible articles by two independent reviewers (among G.P.T, B.K.P, and E.I) to ensure consistency. Reviewers met to discuss the suitability and feasibility of the listed items and to resolve any discrepancies in extraction.

2.4.5 *Synthesis of results*

We extracted and reported study information at the article level, rather than the trial level. This approach was chosen because characteristics and methods may differ across publications from

the same trial program and overlapping participants were of less concern given that the results are not synthesized in this scoping review.

We used descriptive analyses to summarize general characteristics across all included articles and to summarize additional specific methodological characteristics in the subset prioritized for detailed extraction. Summary tables were created for all included articles to present study design, trial phase, allocation, disease, trial internal arm number and sample size. For articles in the prioritized subset, we created detailed tables to summarize reporting characteristics, external control design features, and analytic methods. We did not assess risk of bias in this scoping review as the aim was not to assess the evidence of the included articles but rather to characterize current practices.

2.5 Results

2.5.1 *Selection of pediatric genetic RD externally controlled trials*

From 1,513 articles imported into Covidence for full-text screening, 1,400 were excluded because their analysis did not include an external control (**Figure 1**). Among the remaining 113 articles, 36 were excluded because they did not include participants with a genetic RD and 10 were excluded because they did not make an external control comparison to an efficacy or safety outcome (for example, external data may have been used in a pharmacokinetic model). Sixty-seven articles were ultimately eligible for inclusion. Publication frequency increased slightly between 2014 and 2024 (**eFigure 2**). Of the included articles, 16 (24%) described trials that made a comparison to aggregate external data,[36-51] and 51 (76%) reported a comparison to individual-level external data (**eTable 1**).[52-102] Of these 51, a subset of 18 articles (27%) described trials that attempted to address differences in baseline characteristics between internal

and external arms statistically, beyond adjustment for age and baseline outcome scores, making up the prioritized subset.[52-69] General information was extracted from all 67 articles and additional methodological and reporting information was extracted for the subset of 18 articles reporting on an individual-level external comparison and on trials that used methods to address baseline covariate imbalances.

2.5.2 *Study characteristics of all included externally controlled pediatric genetic RD trials*

Study sites spanned six continents including North America, South America, Europe, Asia, Africa and Oceania (**eFigure 3**). Twenty-three articles (34%) reported on trials with study sites in only one country and 44 (66%) included sites in more than one country. A majority of articles reported on trials including at least one site for their internal arms in the United States (n=54), with many also including sites in the United Kingdom (n=31) or Italy (n=24). Across all articles, 37 different genetic RDs and disease subtypes were represented (**eTable 2**). Duchenne muscular dystrophy (n=17) was the most commonly included disease, followed by spinal muscular atrophy type 1 (n=9) and spinal muscular atrophy type 2 (n = 6). Most articles reported the results of a main trial only (n=41, 61%). Sixteen (24%) reported the results for both a main trial and extension component and 10 (15%) reported the results an extension study exclusively (**Table 1**). Trial phases included phase 1 (n=7), phase 1/2 (n=13), phase 2 (n=21), phase 2/3 (n=4), phase 3 (n=16) and phase 4 (n=1). Thirty-six articles (54%) reported on single-arm studies, 22 (33%) reported two internal arms, and 9 (13%) included three or more. Among the 31 articles reporting on studies with at least two internal arms, 12 (39%) used random allocation. Allocation was to different doses of the investigational treatment, a placebo or standard of care. Sample size across internal arms ranged from 4 to 1,030 participants, with a median of 22. Aside from one larger

study,[53] all articles included fewer than 174 participants across their internal arms. Most articles reported on trials conducted exclusively in pediatric populations (63%, n=42) with the remainder including children and adults.

2.5.3 *Prioritized subset: Externally controlled pediatric genetic RD trials that applied specific analytic methods to address covariate imbalances*

2.5.3.1 *General study and participant characteristics*

Eighteen articles reported on trials that relied on individual-level external control data and applied analytic methods to address covariate imbalance beyond adjusting for age and baseline outcome scores. These articles covered trials of nine different RDs, including cystic fibrosis (n=1), Duchenne muscular dystrophy (n=8), fibrodysplasia ossificans progressiva (n=1), metachromatic leukodystrophy (n=1), mucopolysaccharidosis type 4A (n=1), neuronal ceroid lipofuscinosis type 2 (CLN2) (n=2), spinal muscular atrophy type 1 (n=3), spinal muscular atrophy type 2 (n=2) and spinal muscular atrophy type 3 (n=2) (**eTable 2**).

Ten articles (56%) reported on single-arm trials, while 8 included multiple internal arms (44%), where participants were assigned to receive an intervention, a placebo, or standard of care (**Table 1**). However, only four multi-arm trials compared more than one internal arm to an external control. Among articles with more than one internal arm, some had overlapping participants, but each was compared separately to an external comparator. Seven of the 18 articles reported on a primary trial while 5 reported an extension study and 6 included both (**Table 1**).

Two sets of two articles and one set of three articles were directly linked manuscripts from the same trial program (e.g., main trial report, extension study, interim analysis) and

therefore included the same intervention and completely or largely overlapping participants (**eTable 3**). This included three articles focused on Duchenne muscular dystrophy that evaluated the medication viltolarsen, an antisense oligonucleotide,[57,61,64] two articles focused on Duchenne muscular dystrophy that evaluated the efficacy of eteplirsen, another antisense oligonucleotide,[56,60] and two articles focused on CLN2 that evaluated the efficacy of cerliponase alfa, an enzyme replacement therapy.[55,69] Several other articles were related through shared interventions and/or use of the same external control data source (**eTable 3**). For example, of eight articles reporting on trials for Duchenne muscular dystrophy, all eight relied on the same external control data source to evaluate four different interventions.[56,57,60,61,63,64,67,68]

Apart from articles that included patients with Duchenne muscular dystrophy, which disproportionately affects males, all articles in the prioritized subset reported on trials that included participants of both female and male sex at birth in both their internal and external arms (**Table 2**). However, none reported baseline characteristics or results stratified by sex. Eleven articles (61%) were restricted to pediatric populations, while seven (39%) included both pediatric and adult populations. Age was not reported for external control participants in three articles. Among those that did report age, the range was similar to that of their corresponding internal arm (**Table 3**). Sample sizes ranged from 10 to 516 in internal arms compared to the external control (median = 27) and from 4 to 1,588 participants in external arms (median = 65) (**Table 2**); aside from one larger study,[53] all internal and external arms included fewer than 174 participants (**Table 3**).

2.5.3.2 *Characteristics of external controls*

Use of an external control was planned a priori in 14 of the 18 articles (78%) in the prioritized subset (**Table 2**), however, only two justified their use of this study design in their publication or protocol.[55,62] Thirteen articles made a comparison to one external control arm, three made a comparison to two external control arms, one included three and one included four distinct external control arms (**Table 2; details for each article, Table 3**). All articles in the prioritized subset of 18 articles included external data from at least one observational source, such as a cohort study, natural history study or longitudinal patient registry. Less commonly, external data were derived from previous clinical trials (n=2) or routinely collected data (n=1),[58,65,67] such as health care administrative data or medical chart abstraction. Eight articles reported using historical external data, meaning data were collected for the external comparator prior to the conduct of the trial, one included a concurrent external arm, and five included external data that was partially concurrent (**Tables 2 and 3**). The timing of external data collection was not reported in five articles.

2.5.3.3 *Outcomes and analytic methods, including strategies to address covariate imbalance*

Ten articles (56%) compared at least one primary outcome between internal and external arms, and 17 articles (94%) compared at least one secondary outcome (**Table 4**). Motor function was the most commonly measured outcome, although the measurement tools varied across articles. The outcomes compared between internal and external arms were most often analyzed as a change from baseline (n=16, 89%) or time to event (n=6, 33%). Time axes included age (n=4) and calendar time (n=17). Across all articles, the time at which follow-up began (time zero) for the internal arms was either a designated time within one month of treatment (n=17) or birth

(n=3), depending on the outcome. Twelve articles (66%) did not specify time zero in the external control for at least one outcome. Two sets of three articles (17%) selected birth, or the time at which external control participants were enrolled into their natural history study as time zero for at least one outcome. Two articles (11%) designated time zero as a specific point in time during follow-up in a natural history study.

Methods used to address covariate imbalance across internal and external arms among the 18 articles in the prioritized subset included: individual matching on specified variables (n=8), sibling matching (n=1), group matching on specified variables (n=3), regression analysis conditioning on specific baseline covariates (n=8, one of which used a Bayesian analysis), use of propensity scores with inverse probability of treatment weighting (n=2) and use of propensity score matching (n=2) (**Table 4 and eTable 4)**. Although analytic methods were applied in an effort to address imbalances, only five articles (28%) reported or discussed any balance diagnostics, while the remaining 13 articles (72%) did not discuss the success of their selected method. Furthermore, the variables accounted for using these methods varied widely in comprehensiveness (**eTable 4**).

Statistical analysis methods used to compare outcomes included t-tests, exact binomial tests, analysis of covariance, mixed effects models, negative binomial models, nonparametric maximum likelihood estimates, Kaplan Meier analyses, Cox proportional hazard models and Bayesian compound Poisson models (**Table 4**). Outcome assessment timing was not always comparable between arms. Eight articles (44%) included at least one outcome with differing assessment schedules and four (21%) did not report timing for at least one external outcome. Only one article (5%) reported blinding in the assessment of one of their outcomes for external data, which could include blinding at the level of the outcome assessors or those analyzing the

data. Reporting of missing data was poor, as only three articles (17%) reported missingness in the baseline characteristics of their external arms, one article (6%) reported no missingness and 14 (78%) failed to report missingness. Nine articles (50%) reported missingness in at least one outcome, one study (6%) reported no missingness and eight articles (44%) failed to report any missingness. Only one study (6%) pre-specified how missingness would be addressed specifically in the external arm data. Five articles (28%) reported a sensitivity analysis that included external comparator data, and none performed a quantitative bias analysis with their external comparison data (**Table 4**).

2.6 Discussion

2.6.1 *Summary of evidence*

This scoping review identified 67 articles that reported trials in pediatric patients with genetic RDs that used external data to construct a comparator. The review explored the design, conduct and reporting of these trials. Only 18 articles reported on trials that relied on individual participant data and attempted to address baseline covariate imbalances across arms, beyond adjusting for participant age and baseline score of the outcome, leaving many of these articles susceptible to bias due to residual or unmeasured confounding. Among those that did attempt to adjust for covariate imbalances, the variables considered in the adjustment and the methods applied varied widely in both quantity and complexity. Over half of the 18 articles in the prioritized subset had single internal arms and a majority compared a primary outcome to the external control data, yet reporting on outcome assessment timepoints in the external control data was often poor and the handling of missing data rarely addressed.

Confounding due to differences in baseline characteristics is an important potential source of bias in externally controlled trials, due to lack of random allocation to arms and exacerbated by the use of a different data source, and often a historical data source, as the control arm.[103-105] It is thus concerning that only a minority of externally controlled trials attempted to address baseline covariate imbalance and that those that did frequently considered only a small number of covariates and typically did not evaluate their strategy. The selection of variables and methods may have been subject to the availability of baseline characteristics in the external control data source or constrained by the sample size. Given that RD populations are characteristically small, it may be difficult for authors to consider multiple variables when balancing populations, as the selection of more variables, particularly in matching strategies, may lead to the exclusion of more participants and ultimately reduce an already small sample size. Specific methodological guidance on how to limit confounding between arms in small samples is needed and may be beneficial to other disease areas as well.[21,25]

The potential poor comparability of outcomes across internal and external arms has also been identified as a concern in externally controlled trials.[13,104,106] Some articles in our review failed to report the timing of data collection or time at which follow-up began in the external control data, compared outcomes measured at different or non-equivalent time points across arms, or introduced post-hoc changes to analysis time frames, increasing their susceptibility to bias and further reducing the reliability of their results. For example, two included articles that compared changes from baseline used different time intervals between baseline measurement and the comparison point in the internal and external arms.[52,63] One article initially included additional outcome assessment time points for the internal arm in their analysis, which was later found to introduce bias and removed from the analysis.[66] In addition to this, the definition of

time zero for external controls was often ambiguous and not well reported. As a result, several of the included articles may be at risk of time-related biases, which is another known risk in externally controlled trials.[106-108] For example, several articles compared survival in their internal arms to their corresponding external arms, with age as the underlying time scale for the analysis and birth selected as the start of follow-up ("time zero").[59,62,69,76,92] This can introduce immortal time bias, particularly in time-to-event analyses when an intervention is received some time after birth but participants in the intervention arm must have survived until the time at which treatment was allocated. This may have led to the overestimation of treatment effects in several articles. Sensitivity analyses are an important tool that can be used to assess the robustness of study results and justify the selection of a time scale, time zero and assessment time points, however they were poorly utilized among these articles in this review. Poor uptake and reporting of sensitivity analyses was similarly reported by Liu et al., in their review of externally controlled trials across several disease areas.[21]

Finally, many trialists failed to prespecify how missingness in the external control arm would be addressed and provided limited information on the extent of or reason for missingness in the external comparator arm. If this lack of reporting on missingness is due to the intentional selection of external control participants with little or no missing data, this may introduce selection bias into the results.[109,110] Three other literature reviews, focused on externally controlled trials in other disease-areas similarly highlighted very poor reporting of missingness in external controls across their included articles.[21,24,25]

As a result of these pre-specification, methodological and reporting practices, the validity of the conclusions of many externally controlled trials in pediatric genetic RDs published over the last decade may be at risk. To improve the robustness and rigor of future externally

controlled pediatric genetic RD trials, trialists should prespecify the study design elements in a protocol, ensure that external data sources are appropriate, comprehensive and comparable to the intervention cohort, and apply specific analytic methods to reduce bias and confounding. In addition, transparent reporting and justification of key design elements in the external control, including timing of data collection, definition and comparability of time zero, handling of missing data, and where possible, alignment of outcome assessment time points across study arms, is essential and will ultimately support the reliability and robustness of results. In cases where the alignment of outcome assessment time points across study arms is not possible, as may often be the case when using RWD, these differences should be clearly reported and sensitivity analyses to assess assumptions and residual bias should be conducted. These findings highlight a clear need for comprehensive guidance for the design, methodological conduct and reporting of externally controlled trials pediatric genetic RDs.

2.6.2 *Strengths and limitations*

This scoping review has limitations that should be considered when interpreting our results. First, the initial search strategies of the RareKids-CAN living scoping review were restricted to OVID MEDLINE and the Cochrane CENTRAL, which could have resulted in missing some externally controlled trials. Second, this scoping review was limited by the reporting of the trial authors. Consequently, we were unable to obtain information for all data fields in some articles. To limit missingness, we considered all publicly available documents linked to the study report during extraction, including trial registrations, protocols, statistical analysis plans and supplemental material. Another consequence of our inherent need to rely on author reporting was that certain terms were used inconsistently, for example, the term “matching”. As a result, we

made a set of decision rules that were iterated as needed over time and applied to all trials in the data. Two reviewers were involved to ensure a high degree of consistency. Finally, publication bias favoring positive or statistically significant findings may have influenced the available literature, however, the articles reviewed in this scoping review included both positive and null treatment efficacy results and our interest was in study methods rather than results. Despite these limitations, this scoping review provides important conclusions to inform future externally controlled trials in pediatric genetic RDs.

2.6.3 *Conclusions*

Given the emergence of new disease-modifying therapies such as gene therapies, externally controlled trials in pediatric genetic RDs are increasingly used when RCTs are not possible. However, the methodological and reporting practices of these trials are heterogeneous, and many articles may be subject to important biases related to confounding, time, outcome assessments and missingness. These limitations may compromise the validity of their findings, ultimately impacting regulatory and health technology assessment decision-making and access to treatments. Comprehensive and specific methodological and reporting guidance is needed to support the conduct of robust externally controlled trials in pediatric genetic RDs and to improve access to treatments.

2.7 References for Chapter 2

1. European Commission. Rare diseases: EU research on rare diseases. Accessed April 1, 2025. https://research-and-innovation.ec.europa.eu/research-area/health/rare-diseases_en

2. Canadian Organization for Rare Diseases. About CORD: Key facts. Accessed February 20, 2025. https://www.raredisorders.ca/about/about-cord

3. Orphan Drug Act, Pub. L. No. 97-414, 96 2049-2066 (1983). Accessed April 18, 2025. https://www.govinfo.gov/content/pkg/STATUTE-96/pdf/STATUTE-96-Pg2049.pdf

4. The landscape for rare diseases in 2024. *Lancet Glob Health*. 2024;12(3):e341. doi:10.1016/S2214-109X(24)00056-1

5. Nguengang Wakap S, Lambert DM, Olry A, et al. Estimating cumulative point prevalence of rare diseases: Analysis of the Orphanet database. *Eur J Hum Genet*. 2020;28(2):165-173. doi:10.1038/s41431-019-0508-0

6. Rath A, Salamon V, Peixoto S, et al. A systematic literature review of evidence-based clinical practice for rare diseases: what are the perceived and real barriers for improving the evidence and how can they be overcome? *Trials*. 2017;18(1):556. doi:10.1186/s13063-017-2287-7

7. Baek HJ, Nam Y, Rim YA, Ju JH. Advances in cell and gene therapy for rare disease treatment. *Int J Stem Cells*. 2025;18(4):324-337. doi:10.15283/ijsc25028

8. Jahanshahi M, Gregg K, Davis G, et al. The use of external controls in FDA regulatory decision making. *Ther Innov Regul Sci*. 2021;55(5):1019-1035. doi:10.1007/s43441-021-00302-y

9. Pizzamiglio C, Vernon HJ, Hanna MG, Pitceathly RDS. Designing clinical trials for rare diseases: Unique challenges and opportunities. *Nat Rev Methods Primers*. 2022;2(1):s43586-022-00100-2. doi:10.1038/s43586-022-00100-2

10. Augustine EF, Adams HR, Mink JW. Clinical trials in rare disease: Challenges and opportunities. *J Child Neurol*. 2013;28(9):1142-50. doi:10.1177/0883073813495959

11. United States Food and Drug Administration. *Considerations for the design and conduct of externally controlled trials for drug and biological products guidance for industry: Draft guidance*. 2023. Accessed October 24, 2024. https://www.fda.gov/vaccines-blood-biologics/guidance-compliance-regulatory-information-biologics/biologics-guidances

12. Weberpals J, Wang SV. External controls to study treatment effects in rare diseases: Challenges and future directions. *Clin Pharmacol Ther*. 2024;116(6):1521-1524. doi:10.1002/cpt.3443

13. Sola-Morales O, Curtis LH, Heidt J, et al. Effectively leveraging RWD for external controls: A systematic literature review of regulatory and HTA decisions. *Clin Pharmacol Ther*. 2023;114:325-355. doi:10.1002/cpt.2914

14. Vaghela S, Tanni KA, Banerjee G, Sikirica V. A systematic review of real-world evidence (RWE) supportive of new drug and biologic license application approvals in rare diseases. *Orphanet J Rare Dis*. 2024;19(1):117. doi:10.1186/s13023-024-03111-2

15. Velummailum RR, McKibbon C, Brenner DR, Stringer EA, Ekstrom L, Dron L. Data challenges for externally controlled trials: Viewpoint. *J Med Internet Res*. 2023;25:e43484. doi:10.2196/43484

16. Canada's Drug Agency. *Methods guide for health technology assessment*. 2025. Accessed June 16, 2025. https://www.cda-amc.ca/sites/default/files/MG%20Methods/MG0030-Quantitative-Methods-Manual_Mar%202025.pdf

17. Haute Autorité de Santé. *Methodology for the clinical development of medical devices*. 2021. Accessed June 16, 2025. https://www.has-sante.fr/upload/docs/application/pdf/2021-09/guide_methodology_for_the_clinical_development_of_md.pdf

18. National Insitute for Health and Care Excellence. *NICE real-world evidence framework*. 2022. Accessed June 16, 2025. https://www.nice.org.uk/corporate/ecd9/resources/nice-realworld-evidence-framework-pdf-1124020816837

19. International Council for Harmonization of Technical Requirements for Pharmaceuticals for Human Use. *E8 (R1): General considerations for clinical studies*. 2021. Accessed June 16, 2025. https://database.ich.org/sites/default/files/ICH_E8-R1_Guideline_Step4_2021_1006.pdf

20. International Council for Harmonization of Technical Requirements for Pharmaceuticals for Human Use. *E10: Choice of control group and related issues in clinical trials*. 2000. Accessed June 16, 2025. https://database.ich.org/sites/default/files/E10_Guideline.pdf

21. Liu J, Yao M, Wang M, et al. Design, conduct, and analysis of externally controlled trials. *JAMA Netw Open*. 2025;8(9):e2530277-e2530277. doi:10.1001/jamanetworkopen.2025.30277

22. Goring S, Taylor A, Müller K, et al. Characteristics of non-randomised studies using comparisons with external controls submitted for regulatory approval in the USA and Europe: A systematic review. *BMJ Open*. 2019;9(2):e024895. doi:10.1136/bmjopen-2018-024895

23. Jaksa A, Louder A, Maksymiuk C, et al. A comparison of 7 oncology external control arm case studies: Critiques from regulatory and health technology assessment agencies. *Value Health*. 2022;25(12):1967-1976. doi:10.1016/j.jval.2022.05.016

24. Hermans SJF, van der Maas NG, van Norden Y, et al. Externally controlled studies using real-world data in patients with hematological cancers: A systematic review. *JAMA Oncol*. 2024;10(10):1426-1436. doi:10.1001/jamaoncol.2024.3466

25. Zayadi A, Edge R, Parker CE, et al. Use of external control arms in immune-mediated inflammatory diseases: A systematic review. *BMJ Open*. 2023;13(12):e076677. doi:10.1136/bmjopen-2023-076677

26. Peters MDJ, Marnie C, Tricco AC, et al. Updated methodological guidance for the conduct of scoping reviews. *JBI Evid Implement*. 2021;19(1):3-10. doi:10.1097/xeb.0000000000000277

27. Tricco AC, Lillie E, Zarin W, et al. PRISMA extension for Scoping Reviews (PRISMA-ScR): Checklist and explanation. *Ann Intern Med*. 2018;169(7):467-473. doi:10.7326/m18-0850

28. Pratt Tremblay G, Iverson E, Bliu A, Gravel C, Potter B. Protocol 3: Externally controlled clinical trials for pediatric rare disease: Current practices and guidance. *OSF*. 2025. doi:10.17605/OSF.IO/4MBRN

29. RareKids-CAN: Pediatric rare disease clinical trials and treatment network. Accessed November 1, 2025, https://www.rarekidscan.com

30. Iverson E, Straton C, McGowan J, et al. Protocol 1: Pediatric rare disease clinical trials inventory: A living scoping review protocol. *OSF*. 2025. doi:10.17605/OSF.IO/DRXJY

31. Orphanet. Diseases. Accessed January 20, 2025, https://www.orpha.net/en/disease

32. National Institutes of Health. Genetic and Rare Diseases (GARD) Information Center: Diseases. Accessed January 20, 2025. https://rarediseases.info.nih.gov

33. Veritas Health Innovation. Covidence systematic review software. Accessed February 26, 2025. www.covidence.org

34. Harris PA, Taylor R, Minor BL, et al. The REDCap consortium: Building an international community of software platform partners. *J Biomed Inform*. 2019;95:103208. doi:10.1016/j.jbi.2019.103208

35. Clinical Trials Transformation Initiative (CTTI). Aggregate Analysis of ClinicalTrials.gov (AACT) Database. Accessed July 10, 2025. https://aact.ctti-clinicaltrials.org

36. Auerswald G, Kurnik K, Aledort LM, et al. The EPIC study: A lesson to learn. *Haemophilia*. 2015;21(5):622-628. doi:10.1111/hae.12666

37. Chien Y-H, Lee N-C, Tseng S-H, et al. Efficacy and safety of AAV2 gene therapy in children with aromatic L-amino acid decarboxylase deficiency: An open-label, phase 1/2 trial. *Lancet Child Adolesc Health*. 2017;1(4):265-273. doi:10.1016/S2352-4642(17)30125-6

38. Dvorak CC, Horn BN, Puck JM, et al. A trial of alemtuzumab adjunctive therapy in allogeneic hematopoietic cell transplantation with minimal conditioning for severe combined immunodeficiency. *Pediatr Transplant*. 2014;18(6):609–616. doi:10.1111/petr.12310

39. Gilchrist FJ, Bui S, Gartner S, et al. ALPINE2: Efficacy and safety of 14-day vs 28-day inhaled aztreonam for Pa eradication in children with cystic fibrosis. *J Cyst Fibros*. 2024;23(1):80-86. doi:10.1016/j.jcf.2023.06.008

40. Hankins JS, Aygun B, Nottage K, et al. From infancy to adolescence: Fifteen years of continuous treatment with hydroxyurea in sickle cell anemia. *Medicine*. 2014;93(28):e215. doi:10.1097/MD.0000000000000215

41. Hoppe JE, Kasi AS, Pittman JE, et al. Vanzacaftor-tezacaftor-deutivacaftor for children aged 6-11 years with cystic fibrosis (RIDGELINE Trial VX21-121-105): An analysis from a single-arm, phase 3 trial. *Lancet Respir Med*. 2024;13(3):244-255. doi:10.1016/S2213-2600(24)00407-7

42. Kavakli K, Yang R, Rusen L, Beckmann H, Tseneklidou-Stoeter D, Maas Enriquez M. Prophylaxis vs. on-demand treatment with BAY 81-8973, a full-length plasma protein-free recombinant factor VIII product: Results from a randomized trial (LEOPOLD II). *J Thromb Haemost*. 2015;13(3):360-369. doi:10.1111/jth.12828

43. Lynch DR, Farmer J, Hauser L, et al. Safety, pharmacodynamics, and potential benefit of omaveloxolone in Friedreich ataxia. *Ann Clin Transl Neurol*. 2019;6(1):15-26. doi:10.1002/acn3.660

44. Lynch DR, Hauser L, McCormick A, et al. Randomized, double-blind, placebo-controlled study of interferon-γ 1b in Friedreich Ataxia. *Ann Clin Transl Neurol*. 2019;6(3):546-553. doi:10.1002/acn3.731

45. Manco-Johnson MJ, Bomgaars L, Palascak J, et al. Efficacy and safety of protein C concentrate to treat purpura fulminans and thromboembolic events in severe congenital protein C deficiency. *Thromb Haemost*. 2016;116(1):58-68. doi:10.1160/th15-10-0786

46. Mendell JR, Al-Zaidy S, Shell R, et al. Single-dose gene-replacement therapy for spinal muscular atrophy. *N Engl J Med*. 2017;377(18):1713-1722. doi:10.1056/NEJMoa1706198

47. Mendell JR, Sahenk Z, Lehman K, et al. Assessment of systemic delivery of rAAVrh74.MHCK7.micro-dystrophin in children with Duchenne muscular dystrophy: A nonrandomized controlled trial. *JAMA Neurol*. 2020;77(9):1122-1131. doi:10.1001/jamaneurol.2020.1484

48. Patterson M, Ramaswami U, Donald A, et al. Disease-modifying, neuroprotective effect of N-acetyl-L-leucine in adult and pediatric patients with Niemann-Pick disease type C. *Neurology*. 2024;105(1):e213589. doi:10.1101/2024.10.11.24315318

49. Saxena K, Lalezari S, Oldenburg J, et al. Efficacy and safety of BAY 81-8973, a full-length recombinant factor VIII: Results from the LEOPOLD I trial. *Haemophilia*. 2016;22(5):706-712. doi:10.1111/hae.12952

50. Taylor M, Jefferies J, Byrne B, et al. Cardiac and skeletal muscle effects in the randomized HOPE-Duchenne trial. *Neurology*. 2019;92(8):e866-e878. doi:10.1212/WNL.0000000000006950

51. Windyga J, Lissitchkov T, Stasyshyn O, et al. Pharmacokinetics, efficacy and safety of BAX326, a novel recombinant factor IX: A prospective, controlled, multicentre phase I/III study

in previously treated patients with severe (FIX level <1%) or moderately severe (FIX level ≤2%) haemophilia B. *Haemophilia*. 2014;20(1):15-24. doi:10.1111/hae.12228

52. Hendriksz CJ, Parini R, AlSayed MD, et al. Long-term endurance and safety of elosulfase alfa enzyme replacement therapy in patients with Morquio A syndrome. *Mol Genet Metab*. 2016;119(1-2):131-143. doi:10.1016/j.ymgme.2016.05.018

53. Konstan MW, McKone EF, Moss RB, et al. Assessment of safety and efficacy of long-term treatment with combination lumacaftor and ivacaftor therapy in patients with cystic fibrosis homozygous for the F508del-CFTR mutation (PROGRESS): A phase 3, extension study. *Lancet Respir Med*. 2017;5(2):107-118. doi:10.1016/S2213-2600(16)30427-1

54. Krosschell KJ, Kissel JT, Townsend EL, et al. Clinical trial of L-carnitine and valproic acid in spinal muscular atrophy type I. *Muscle Nerve*. 2018;57(2):193-199. doi:10.1002/mus.25776

55. Schulz A, Ajayi T, Specchio N, et al. Study of intraventricular cerliponase alfa for CLN2 disease. *N Engl J Med*. 2018;378(20):1898-1907. doi:10.1056/NEJMoa1712649

56. Khan N, Eliopoulos H, Han L, et al. Eteplirsen treatment attenuates respiratory decline in ambulatory and non-ambulatory patients with Duchenne muscular dystrophy. *J Neuromuscul Dis*. 2019;6(2):213-225. doi:10.3233/JND-180351

57. Clemens PR, Rao VK, Connolly AM, et al. Safety, tolerability, and efficacy of viltolarsen in boys with Duchenne muscular dystrophy amenable to exon 53 skipping: A phase 2 randomized clinical trial. *JAMA Neurol*. 2020;77(8):982-991. doi:10.1001/jamaneurol.2020.1264

58. Muntoni F, Bertini E, Comi G, et al. Long-term follow-up of patients with type 2 and non-ambulant type 3 spinal muscular atrophy (SMA) treated with olesoxime in the OLEOS trial. *Neuromuscul Disord*. 2020;30(12):959-969. doi:10.1016/j.nmd.2020.10.008

59. Day JW, Finkel RS, Chiriboga CA, et al. Onasemnogene abeparvovec gene therapy for symptomatic infantile-onset spinal muscular atrophy in patients with two copies of SMN2 (STR1VE): An open-label, single-arm, multicentre, phase 3 trial. *Lancet Neurol*. 2021;20(4):4284-4293. doi:10.1016/S1474-4422(21)00001-6

60. McDonald CM, Shieh PB, Abdel-Hamid HZ, et al. Open-label evaluation of eteplirsen in patients with Duchenne muscular dystrophy amenable to exon 51 skipping: PROMOVI trial. *J Neuromuscul Dis*. 2021;8(6):989-1001. doi:10.3233/JND-210643

61. Clemens PR, Rao VK, Connolly AM, et al. Long-term functional efficacy and safety of viltolarsen in patients with Duchenne muscular dystrophy. *J Neuromuscul Dis*. 2022;9(4):493-501. doi:10.3233/JND-220811

62. Fumagalli F, Calbi V, Natali Sora MG, et al. Lentiviral haematopoietic stem-cell gene therapy for early-onset metachromatic leukodystrophy: Long-term results from a non-randomised, open-label, phase 1/2 trial and expanded access. *Lancet*. 2022;399(10322):372-383. doi:10.1016/S0140-6736(21)02017-1

63. Mah JK, Clemens PR, Guglieri M, et al. Efficacy and safety of vamorolone in Duchenne muscular dystrophy: A 30-month nonrandomized controlled open-label extension trial. *JAMA Netw Open*. 2022;5(5):e2144178. doi:10.1001/jamanetworkopen.2021.44178

64. Clemens PR, Rao VK, Connolly AM, et al. Efficacy and safety of viltolarsen in boys with Duchenne muscular dystrophy: Results from the phase 2, open-label, 4-year extension study. *J Neuromuscul Dis*. 2023;10(3):439-447. doi:10.3233/JND-221656

65. Mercuri E, Baranello G, Boespflug-Tanguy O, et al. Risdiplam in types 2 and 3 spinal muscular atrophy: A randomised, placebo-controlled, dose-finding trial followed by 24 months of treatment. *Eur J Neurol*. 2023;30(7):1945-1956. doi:10.1111/ene.15499

66. Pignolo RJ, Hsiao EC, Al Mukaddam M, et al. Reduction of new heterotopic ossification (HO) in the open-label, phase 3 MOVE trial of palovarotene for fibrodysplasia ossificans progressiva (FOP). *J Bone Miner Res*. 2023;38(3):381-394. doi:10.1002/jbmr.4762

67. Zaidman CM, Proud CM, McDonald CM, et al. Delandistrogene moxeparvovec gene therapy in ambulatory patients (aged ≥4 to <8 Years) with Duchenne muscular dystrophy: 1-year interim results from study SRP-9001-103 (ENDEAVOR). *Ann Neurol*. 2023;94(5):955-968. doi:10.1002/ana.26755

68. Harper AD, Topaloglu H, Mercuri E, et al. Safety and efficacy of viltolarsen in ambulatory and nonambulatory males with Duchenne muscular dystrophy. *Sci Rep*. 2024;14(1):23488. doi:10.1038/s41598-024-70783-y

69. Schulz A, Specchio N, de Los Reyes E, et al. Safety and efficacy of cerliponase alfa in children with neuronal ceroid lipofuscinosis type 2 (CLN2 disease): An open-label extension study. *Lancet Neurol*. 2024;23(1):60-70. doi:10.1016/S1474-4422(23)00384-8

70. Hacein-Bey-Abina S, Pai S-Y, Gaspar HB, et al. A modified gamma-retrovirus vector for X-linked severe combined immunodeficiency. *N Engl J Med*. 2014;371(15):1407-1417. doi:10.1056/NEJMoa1404588

71. Levin SW, Baker EH, Zein WM, et al. Oral cysteamine bitartrate and N-acetylcysteine for patients with infantile neuronal ceroid lipofuscinosis: A pilot study. *Lancet Neurol*. 2014;13(8):777-787. doi:10.1016/S1474-4422(14)70142-5

72. Jones SA, Bialer M, Parini R, et al. Safety and clinical activity of elosulfase alfa in pediatric patients with Morquio A syndrome (mucopolysaccharidosis IVA) less than 5 y. *Pediatr Res*. 2015;78(6):717-722. doi:10.1038/pr.2015.169

73. Finkel RS, Chiriboga CA, Vajsar J, et al. Treatment of infantile-onset spinal muscular atrophy with nusinersen: A phase 2, open-label, dose-escalation study. *Lancet*. 2016;388(10063):3017-3026. doi:10.1016/S0140-6736(16)31408-8

74. Whyte MP, Madson KL, Phillips D, et al. Asfotase alfa therapy for children with hypophosphatasia. *JCI insight*. 2016;1(9):e85971. doi:10.1172/jci.insight.85971

75. Whyte MP, Rockman-Greenberg C, Ozono K, et al. Asfotase alfa treatment improves survival for perinatal and infantile hypophosphatasia. *J Clin Endocrinol Metab*. 2016;101(1):334-342. doi:10.1210/jc.2015-3462

76. Jones SA, Rojas-Caro S, Quinn AG, et al. Survival in infants treated with sebelipase Alfa for lysosomal acid lipase deficiency: An open-label, multicenter, dose-escalation study. *Orphanet J Rare Dis*. 2017;12(1):25. doi:10.1186/s13023-017-0587-3

77. Ory DS, Ottinger EA, Farhat NY, et al. Intrathecal 2-hydroxypropyl-beta-cyclodextrin decreases neurological disease progression in Niemann-Pick disease, type C1: A non-randomised, open-label, phase 1-2 trial. *Lancet*. 2017;390(10104):1758-1768. doi:10.1016/S0140-6736(17)31465-4

78. Tardieu M, Zerah M, Gougeon M-L, et al. Intracerebral gene therapy in children with mucopolysaccharidosis type IIIB syndrome: An uncontrolled phase 1/2 clinical trial. *Lancet Neurol*. 2017;16(9):712-720. doi:10.1016/S1474-4422(17)30169-2

79. Kinane TB, Mayer OH, Duda PW, Lowes LP, Moody SL, Mendell JR. Long-term pulmonary function in Duchenne muscular dystrophy: Comparison of eteplirsen-treated patients to natural history. *J Neuromuscul Dis*. 2018;5(1):47-58. doi:10.3233/JND-170272

80. Connolly AM, Zaidman CM, Golumbek PT, et al. Twice-weekly glucocorticosteroids in infants and young boys with Duchenne muscular dystrophy. *Muscle Nerve*. 2019;59(6):650-657. doi:10.1002/mus.26441

81. De Vivo DC, Bertini E, Swoboda KJ, et al. Nusinersen initiated in infants during the presymptomatic stage of spinal muscular atrophy: Interim efficacy and safety results from the Phase 2 NURTURE study. *Neuromuscul Disord*. 2019;29(11):842-856. doi:10.1016/j.nmd.2019.09.007

82. Hoffman EP, Schwartz BD, Mengle-Gaw LJ, et al. Vamorolone trial in Duchenne muscular dystrophy shows dose-related improvement of muscle function. *Neurology*. 2019;93(13):e1312-e1323. doi:10.1212/WNL.0000000000008168

83. Polgreen LE, Lund TC, Braunlin E, et al. Clinical trial of laronidase in Hurler syndrome after hematopoietic cell transplantation. *Pediatr Res*. 2020;87(1):104-111. doi:10.1038/s41390-019-0541-2

84. Smith EC, Conklin LS, Hoffman EP, et al. Efficacy and safety of vamorolone in Duchenne muscular dystrophy: An 18-month interim analysis of a non-randomized open-label extension study. *PLoS Med*. 2020;17(9):e1003222. doi:10.1371/journal.pmed.1003222

85. Sondhi D, Kaminsky SM, Hackett NR, et al. Slowing late infantile Batten disease by direct brain parenchymal administration of a rh.10 adeno-associated virus expressing CLN2. *Sci Transl Med*. 2020;12(572):eabb5413. doi:10.1126/scitranslmed.abb5413

86. Wagner KR, Abdel-Hamid HZ, Mah JK, et al. Randomized phase 2 trial and open-label extension of domagrozumab in Duchenne muscular dystrophy. *Neuromuscul Disord*. 2020;30(6):492-502. doi:10.1016/j.nmd.2020.05.002

87. Darras BT, Masson R, Mazurkiewicz-Beldzinska M, et al. Risdiplam-treated infants with type 1 spinal muscular atrophy versus historical controls. *N Engl J Med*. 2021;385(5):427-435. doi:10.1056/NEJMoa2102047

88. Deiva K, Ausseil J, de Bournonville S, et al. Intracerebral gene therapy in four children with Sanfilippo B syndrome: 5.5-year follow-up results. *Hum Gene Ther*. 2021;32(19-20):1251-1259. doi:10.1089/hum.2021.135

89. Gentner B, Tucci F, Galimberti S, et al. Hematopoietic stem- and progenitor-cell gene therapy for Hurler syndrome. *N Engl J Med*. 2021;385(21):1929-1940. doi:10.1056/NEJMoa2106596

90. Ghosh A, Rust S, Langford-Smith K, et al. High dose genistein in Sanfilippo syndrome: A randomised controlled trial. *J Inherit Metab Dis*. 2021;44(5):1248-1262. doi:10.1002/jimd.12407

91. Tsabari R, Simchovitz E, Lavi E, et al. Safety and clinical outcome of tamoxifen in Duchenne muscular dystrophy. *Neuromuscul Disord*. 2021;31(9):803-813. doi:10.1016/j.nmd.2021.05.005

92. Vijay S, Brassier A, Ghosh A, et al. Long-term survival with sebelipase alfa enzyme replacement therapy in infants with rapidly progressive lysosomal acid lipase deficiency: Final results from 2 open-label studies. *Orphanet J Rare Dis*. 2021;16(1):13. doi:10.1186/s13023-020-01577-4

93. Kato M, Kada A, Shiraishi H, et al. Sirolimus for epileptic seizures associated with focal cortical dysplasia type II. *Ann Clin Transl Neurol*. 2022;9(2):181-192. doi:10.1002/acn3.51505

94. Magnani A, Semeraro M, Adam F, et al. Long-term safety and efficacy of lentiviral hematopoietic stem/progenitor cell gene therapy for Wiskott-Aldrich syndrome. *Nat Med*. 2022;28(1):71-80. doi:10.1038/s41591-021-01641-x

95. Strauss KA, Farrar MA, Muntoni F, et al. Onasemnogene abeparvovec for presymptomatic infants with two copies of SMN2 at risk for spinal muscular atrophy type 1: The phase III SPR1NT trial. *Nat Med*. 2022;28(7):1381-1389. doi:10.1038/s41591-022-01866-4

96. Strauss KA, Farrar MA, Muntoni F, et al. Onasemnogene abeparvovec for presymptomatic infants with three copies of SMN2 at risk for spinal muscular atrophy: The phase III SPR1NT trial. *Nat Med*. 2022;28(7):1390-1397. doi:10.1038/s41591-022-01867-3

97. Crawford TO, Swoboda KJ, De Vivo DC, et al. Continued benefit of nusinersen initiated in the presymptomatic stage of spinal muscular atrophy: 5-year update of the NURTURE study. *Muscle Nerve*. 2023;68(2):157-170. doi:10.1002/mus.27853

98. Henzi BC, Schmidt S, Nagy S, et al. Safety and efficacy of tamoxifen in boys with Duchenne muscular dystrophy (TAMDMD): A multicentre, randomised, double-blind, placebo-controlled, phase 3 trial. *Lancet Neurol*. 2023;22(10):890-899. doi:10.1016/S1474-4422(23)00285-5

99. Seo J-H, Kosuga M, Hamazaki T, Shintaku H, Okuyama T. Intracerebroventricular enzyme replacement therapy in patients with neuronopathic mucopolysaccharidosis type II: Final report of 5-year results from a Japanese open-label phase 1/2 study. *Mol Genet Metab*. 2023;140(4):107709. doi:10.1016/j.ymgme.2023.107709

100. Shieh PB, Kuntz NL, Dowling JJ, et al. Safety and efficacy of gene replacement therapy for X-linked myotubular myopathy (ASPIRO): A multinational, open-label, dose-escalation trial. *Lancet Neurol*. 2023;22(12):1125-1139. doi:10.1016/s1474-4422(23)00313-7

101. Chiriboga CA, Bruno C, Duong T, et al. JEWELFISH: 24-month results from an open-label study in non-treatment-naive patients with SMA receiving treatment with risdiplam. *J Neurol*. 2024;271(8):4871-4884. doi:10.1007/s00415-024-12318-z

102. Sondhi D, Kaminsky SM, Rosenberg JB, Rostami MR, Hackett NR, Crystal RG. Twenty-year survival analysis of adeno-associated virus vector serotype 2-mediated gene therapy to the central nervous system for CLN2 disease. *Hum Gene Ther*. 2024;36(1-2):28-35. doi:10.1089/hum.2024.182

103. Nuño MM, Pugh SL, Ji L, Piao J, Dignam JJ, Steingrimsson JA. On the use of external controls in clinical trials. *J Natl Cancer Inst Monogr*. 2025;2025(68):30-34. doi:10.1093/jncimonographs/lgae046

104. Seeger JD, Davis KJ, Iannacone MR, et al. Methods for external control groups for single arm trials or long-term uncontrolled extensions to randomized clinical trials. *Pharmacoepidemiol Drug Saf*. 2020;29(11):1382-1392. doi:10.1002/pds.5141

105. Schmidli H, Häring DA, Thomas M, Cassidy A, Weber S, Bretz F. Beyond randomized clinical trials: Use of external controls. *Clin Pharmacol Ther*. 2020;107(4):806-816. doi:10.1002/cpt.1723

106. Liu J, Barrett JS, Leonardi ET, et al. Natural history and real-world data in rare diseases: Applications, limitations, and future perspectives. *J Clin Pharmacol*. 2022;62(Suppl 2):S38-s55. doi:10.1002/jcph.2134

107. Suissa S. Single-arm trials with historical controls: Study designs to avoid time-related biases. *Epidemiology*. 2021;32(1):94-100. doi:10.1097/ede.0000000000001267

108. Burcu M, Dreyer NA, Franklin JM, et al. Real-world evidence to support regulatory decision-making for medicines: Considerations for external control arms. *Pharmacoepidemiol Drug Saf*. Oct 2020;29(10):1228-1235. doi:10.1002/pds.4975

109. Peskoe SB, Arterburn D, Coleman KJ, Herrinton LJ, Daniels MJ, Haneuse S. Adjusting for selection bias due to missing data in electronic health records-based research. *Stat Methods Med Res*. 2021;30(10):2221-2238. doi:10.1177/09622802211027601

110. Hernán MA, Hernández-Díaz S, Robins JM. A structural approach to selection bias. *Epidemiology*. 2004;15(5):615-625. doi:10.1097/01.ede.0000135174.63482.43

2.8 Tables and Figures

Table 1. General study characteristics of the trials reported in the articles included in the scoping review, overall and by subgroup

Characteristic	Overall[1] (N=67)	Other[a] (n=49)	Prioritized subset[a,b] (n=18)
Type of data for external comparison			
Aggregate	16 (24%)	16 (33%)	0 (0%)
Individual	51 (76%)	32 (67%)	18 (100%)
Pilot or feasibility study			
Main trial	62 (93%)	45 (92%)	17 (94%)
Pilot or feasibility trial	5 (7%)	4 (8%)	1 (6%)
Both	0 (0%)	0 (0%)	0 (0%)
Type of study			
Main trial only	41 (61%)	34 (69%)	7 (39%)
Extension only	10 (15%)	5 (10%)	5 (28%)
Both	16 (24%)	10 (20%)	6 (33%)
Trial phase			
Phase I	7 (11%)	6 (14%)	1 (6%)
Phase I/II	13 (21%)	9 (21%)	4 (22%)
Phase II	21 (34%)	13 (30%)	8 (44%)
Phase II/III	4 (6%)	4 (8%)	0 (0%)
Phase III	16 (26%)	11 (25%)	5 (28%)
Phase IV	1 (2%)	1 (2%)	0 (0%)
Not reported	5	5	0
Number of internal arms			
1	36 (54%)	26 (53%)	10 (56%)
2	22 (33%)	18 (37%)	4 (22%)
≥3	9 (13%)	5 (10%)	4 (22%)
Allocation of internal arms[c]			
Randomized	12 (39%)	8 (35%)	4 (50%)
Non-randomized	19 (61%)	15 (65%)	4 (50%)
Not applicable	36	26	10
Sample size across all study internal arm(s)	22 (4, 1,030)	18 (4, 174)	34 (16, 1,030)
Internal arm participants age groups			
Children only (<18)	42 (63%)	31 (63%)	11 (61%)
Children and adults	25 (37%)	18 (37%)	7 (39%)

[a]n (%) for categorical variables; Median (Min, Max) for continuous variables

[b]Prioritized subset: used individual-level data for external control group(s) and reported using an analytic method to address baseline covariate imbalances beyond adjusting for age and baseline score of outcome.

[c]Refers to the allocation across internal arms for trials with two or more internal arms.

Table 2. General characteristics and characteristics of external control data for prioritized subset of 18 articles

Characteristic	Prioritized Subset[a,b] (n=18)
Pre-specification of external control	
Yes	14 (78%)
No	2 (11%)
Not reported	2 (11%)
Rationale provided for selection of external control design	
Yes	2 (11%)
No	16 (89%)
Number of external arms	
1	13 (72%)
2	3 (17%)
≥3	2 (11%)
Sample size across internal arms compared to external arms	24 (10, 516)
Sample size across external arms	65 (4, 1,588)
Sex assigned at birth included	
Female	0
Male	8 (44%)
Both female and male	10 (56%)
Data source[c]	
Observational cohort	23
Clinical trial	2
Routinely collected	1
Timing of external control[c]	
Concurrent	1
Historical	8
Partially concurrent	8
Not reported	4

[a]n (%) for categorical variables; Median (Min, Max) for continuous variables

[b]Prioritized subset: used individual-level data for external control group(s) and reported using an analytic method to address baseline covariate imbalances beyond adjusting for age and baseline score of outcome.

[c]More than one answer was possible per study

Table 3. Characteristics of internal arms compared to external controls in studies that used individual-level external data and attempted to address differences in baseline covariates beyond adjustment for age and baseline outcome scores (n=18)

		Internal arm(s) compared to external arm(s)				External arm(s)				
Article	**Rare disease**	**Intervention**	**Sample size**	**Sex at birth**	**Age (years)**	**Data source type**	**Type of control**	**Sample size**	**Sex at birth**	**Age (years)**
(Konstan et al., 2017)[53]	CF	Lumacaftor and Ivacaftor	516	F/M	25.8[b]	Observational cohort	Partial	1588	F/M	NR
(Clemens et al., 2020)[57]	DMD	Viltolarsen	16	M	4-10	Observational cohort	Historical	65	M	4-10
(Clemens et al., 2022)[61]	DMD	Viltolarsen	16	M	4-10	Observational cohort	Historical	65	M	4-10
(Clemens et al., 2023)[64]	DMD	Viltolarsen	16	M	4-10	Observational cohort	Historical	65	M	4-10
(Harper et al., 2024)[68]	DMD	Viltolarsen	A1 = 10 A2 = 10	M	A1 = 8-12 A2 = 8-26	Observational cohort	Historical	48	M	8-21
(Khan et al., 2019)[56]	DMD	Eteplirsen	A1 = 12 A2 = 20 A3 = 42	M	A1 = 10-11 A2 = 10-18 A3 = 10-17	Observational cohort	NR	A1 = 172 A2 = 20 A3 = 148	M	A1 = 10-18 A2 = 10-18 A3 = 10-18
(McDonald et al., 2021)[60]	DMD	Eteplirsen	79	M	7-16	Observational cohort	NR	A1 = 20 A2 = 20	M	A1 = 10-18 A2 = NR
(Mah et al., 2022)[63]	DMD	Vamorolone	23	M	4-8	Observational cohort	A1: Historical A2: Partial	A1 = 75 A2 = 110	M	A1 = 4-8 A2[b] = 6.0
(Zaidman et al., 2023)[67]	DMD	Delandistrogene moxeparvovec	20	M	4-8	Observational cohort and clinical trial	NR	91	M	4-7
(Pignolo et al., 2023)[66]	FOP	Palovarotene	99	F/M	4-61	Observational cohort	NR	111	F/M	4-56
(Fumagalli et al., 2022)[62]	MLD	Atidarsagene autotemcel	A1 = 16 A2 = 13	F/M	A1[b] = 1.1 A2[b] = 5.5	Observational cohort	A1: Partial A2: Partial A3: Partial A4: Partial	A1 = 19 A2 = 12 A3[a] = 7 A4[a] = 4	F/M	A1[b] = 1.7 A2[b] = 4.3 A3: NR A4: NR

(Hendriksz et al., 2016)[52]	MPS IVA	Elosulfase alfa	A1 = 173 A2 = 124 A3 = 43	F/M	A1 = 5-58 A2 = 5-50 A3 = 5-42	Observational cohort	A1: NR A2: NR	A1 = 97 A2 = 79	F/M	A1 = 5-65 A2 = 5-65
(Schulz et al., 2018)[55]	CLN2	Cerliponase alfa	24	F/M	3-9	Observational cohort	Historical	42	F/M	NR
(Schulz et al., 2024)[69]	CLN2	Cerliponase alfa	24	F/M	3-9	Observational cohort	Historical	42	F/M	NR
(Krosschell et al., 2018)[54]	SMA type 1	Levocarnitine and Valproic acid	39	F/M	0-1	Observational cohort	Partial	57	F/M	0-1
(Day et al., 2021)[59]	SMA type 1	Onasemnogene abeparvovec	22	F/M	0-1	Observational cohort	Historical	23	F/M	0-15
(Mercuri et al., 2023)[65]	SMA type 2 and 3	Risdiplam	51	F/M	2-24	Observational cohort and clinical trial	Partial	78	F/M	2-28
(Muntoni et al., 2020)[58]	SMA type 2 and 3	Olesoxime	131	F/M	8-34	Observational cohort and routinely collected	Concurrent	71	F/M	6-39

Note. Rare diseases: CF: Cystic Fibrosis, DMD: Duchenne muscular dystrophy, FOP: Fibrodysplasia ossificans progressiva, MLD: Metachromatic leukodystrophy, SMA: Spinal muscular atrophy, MPS: Mucopolysaccharidosis and CLN2: Neuronal ceroid lipofuscinosis type 2. A: Arm, F: female, M: male, Partial: partially concurrent, NR: not reported.
[a]Sample size was only reported after matching.
[b]Only reported as a mean in article.

Table 4. Analytic and reporting characteristics of studies that used individual-level external data and attempted to address differences in baseline covariates beyond adjustment for age and baseline outcome scores (n=18)

Characteristic	Prioritized Subset[a,b] (n=18)
Type of outcome(s) compared to external control arm[c]	
Any primary outcomes	10 (56%)
Any secondary or exploratory outcomes	17 (94%)
Metric of outcome(s) compared to external control arm[c]	
Change from baseline	16 (89%)
Time-to-event	6 (33%)
Value at a fixed time point	1 (6%)
Time scale[c]	
Age	4 (22%)
Calendar time	17 (94%)
Time zero in internal arm(s)[c]	
Birth	3 (17%)
Within one month of treatment initiation	17 (94%)
Time zero in external arm(s)[c]	
Birth	3 (17%)
Enrollment into external control source	3 (17%)
Specific time point during follow-up in natural history study	2 (11%)
Unspecified	12(66%)
Assessment time points were the same between arms for all outcomes[d]	
Yes	10 (56%)
No	8 (44%)
Did not report the assessment time point(s) for ≥1 outcome	4 (22%)
Blinding in the external control arm for any outcome assessment or analysis: Yes	1 (5%)
Analytic method used to address baseline differences[c]	
Individual matching	8 (44%)
Regression analysis conditioning of specific covariates	8 (44%)
Group matching	3 (17%)
Propensity score: inverse probability of treatment weighting	2 (11%)
Propensity score: matching	2 (11%)
Sibling matching	1 (6%)
Balancing success discussed or evaluated: Yes	3 (17%)
Statistical method[c]	
Mixed model for repeated measures	7 (39%)
Kaplan Meier survival analysis	4 (22%)
Analysis of covariance	3 (17%)
Linear mixed effects model	3 (17%)
T-test	3 (17%)
Cox proportional hazards model	2 (11%)
Bayesian compound Poisson model	1 (6%)

Characteristic	Prioritized Subset[a,b] (n=18)
Logistic regression	1 (6%)
Negative binomial model	1 (6%)
Nonparametric maximum likelihood estimation	1 (6%)
Non-linear mixed effects model	1 (6%)
One-sided exact binomial test	1 (6%)
Missingness in baseline covariates of external arm	
Yes	3 (17%)
No	1 (6%)
Not reported	14 (78%)
Missingness in outcome measures of external arm	
Yes	9 (50%)
No	1 (6%)
Not reported	8 (44%)
Authors prespecified how missingness in external arms would be addressed	
Yes	1 (6%)
No	17 (94%)
Reported a sensitivity analysis that included external data	
Yes	5 (28%)
No	13 (72%)
Reported a quantitative bias analysis that included external data	
Yes	0
No	18 (100%)

[a]n (%)

[b]Prioritized subset: used individual-level data for external control group(s) and reported using an analytic method to address baseline covariate imbalances beyond adjusting for age and baseline score of outcome.

[c]More than one answer was possible per article.

[d]Apart from baseline outcome assessment.

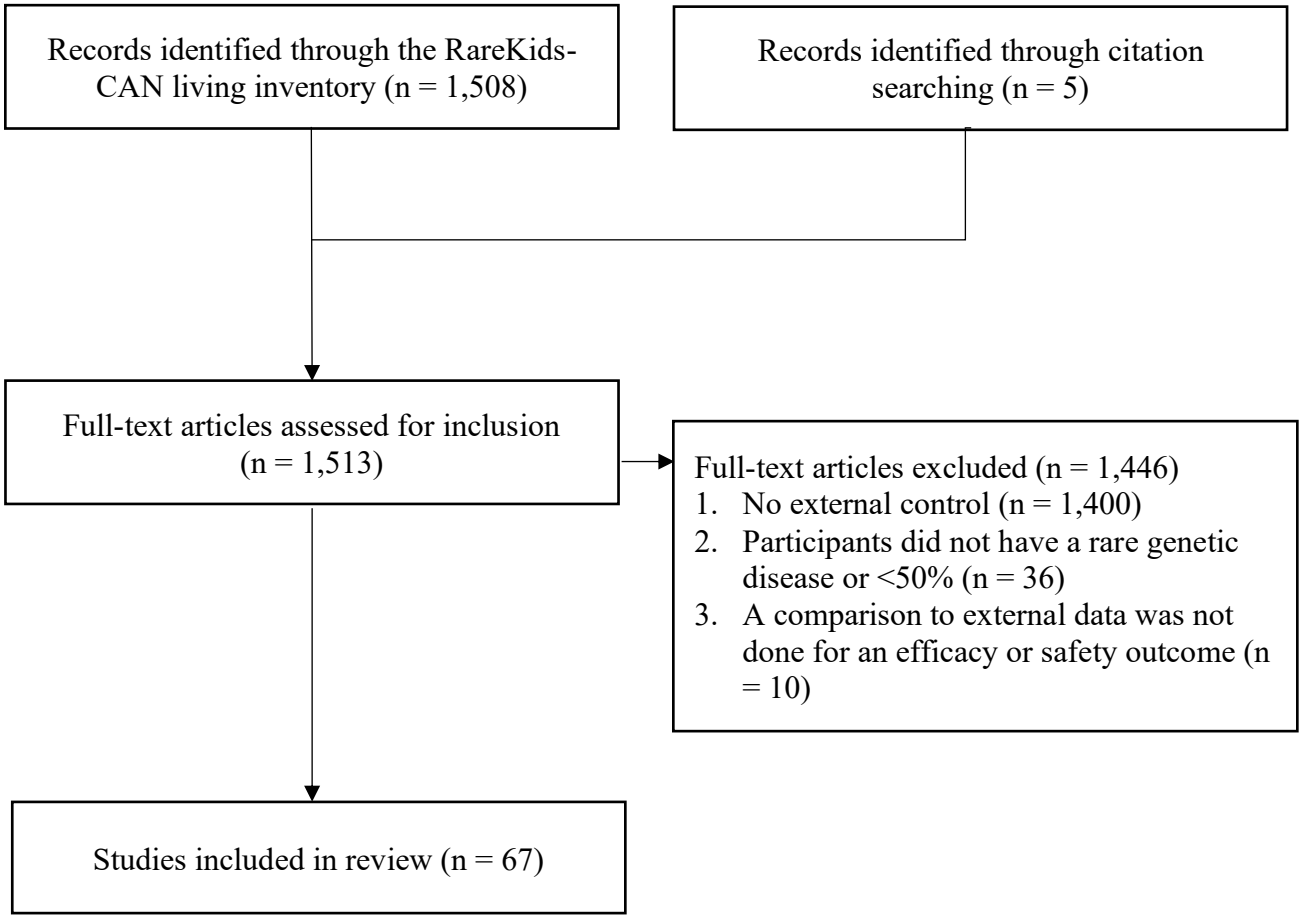


Figure 1. PRISMA flow diagram depicting the identification, screening and selection of articles included in the scoping review

Note. Reasons for exclusion are listed in order of exclusion. The number of studies included in the RareKids-CAN living inventory is up to date as of August 12th, 2026.

Supplementary Online Content

eMethods.

1. Full search strategy used for the RareKids-CAN living scoping review
2. Data items extracted from all articles that met the initial inclusion criteria (N=67)
3. Data items extracted from the subset of articles that met the additional inclusion criteria for detailed extraction (n=18)

eTable 1. Articles included in the scoping review

eTable 2. Genetic RDs represented across all included articles

eTable 3. Related externally controlled trials that attempted to address differences in baseline covariates between arms

eTable 4. Analytic methods used to address imbalances across the internal and external arms and the covariates considered

eFigure 1. PRISMA flow diagram depicting the identification, screening and selection of articles included in the RareKids-CAN living scoping review

eFigure 2. Publication year of included articles

eFigure 3. Global distribution of overall included studies' sites

Data sharing statement: All data reported is publicly available and included in the evidence tables found in the main articles or appendices.

eMethods 1. Final search strategy used for the RareKids-CAN living scoping review by database

Database: Ovid MEDLINE(R) ALL <1946 to February 26, 2025>

---

1 fragile X Syndrome/ or ((fraxa or "fra(x)" or frax or fraxe or FXS or marker x or martin bell) adj syndrome*).tw,kf. or FRAXA.tw,kf. or (fragile X adj3 syndrome*).tw,kf. (8065)

2 Fabry disease/ or (fabry* disease* or alpha galactosidase a deficienc* or anderson fabry diseas* or angiokeratoma* or GLA deficienc* or hereditary dystopic lipidos* or ceramide trihexosidase deficienc*).tw,kf. (6484)

3 Cystic Fibrosis/ or (cystic fibrosis or fibrocystic disease* or mucoviscidosis or pancreas* fibrosis or pancrea* fibrocystic disease*).tw,kf. (60776)

4 muscular atrophy, spinal/ or (spinal muscular atroph* or spinal amyotroph* or SMA type 1 or SMA type I or SMA-I or SMA1 or SMA type 2 or SMA type II or SMA-II or SMA2 or SMA type 3 or SMA type III or SMA-III or SMA3 or "SMN related SMA" or chromosome 5 SMA or neurogenic scapuloperoneal amyotroph* or progressive muscular atroph* or myelopathic atroph* or bulbospinal neuronopath* or hereditary motor neuronopath* or myelopathic muscular atroph*).tw,kf. or ((dubowitz or Werdnig Hoffmann or kugelberg welander) adj disease*).tw,kf. (10167)

5 exp Mucopolysaccharidoses/ or (mucopolysaccharidos* or MPS disorder* or lipochondrodystroph* or polydystrophic dwarfism).tw,kf. or ((alpha-L-iduronidase or Iduronate 2-sulfatase or i2s or Hyaluronidase or Beta-glucuronidase or ARSB or ASB or Arylsulfatase B or N-acetylgalactosamine 4-sulfatase or gusb) adj deficienc*).tw,kf. or ((hurler* or pfaundler-hurler or scheie or hurler-scheie or hunter* or Sanfilippo or Morquio or Sly or Maroteaux-Lamy or Maroteaux) adj2 (diseas* or syndrome*)).tw,kf. (9433)

6 Gaucher Disease/ or (non-neuronopathic GD or neuronopathic GD or Glucocerebrosidase deficienc* or Glucosylceramidase deficienc* or Acid beta glucosidase deficienc* or glucosyl cerebroside lipidos#s or glucosylceramide beta glucosidase deficienc* or glucosylceramide lipidos#s or kerasin histiocytos#s or lipoidos#s histiocytos* or kerasin thesaurismos#s).tw,kf. or ((gaucher* or Gaucher-like) adj (diseas* or syndrome*)).tw,kf. (6621)

7 neuronal Ceroid-Lipofuscinoses/ or (neuronal ceroid-lipofuscinos* or neuronal ceroid lipofuscinos* or cerebroretinal degeneration or parkinsonism-neuronal ceroid lipofuscinos* or parkinsonism neuronal ceroid lipofuscinos* or turkish variant late infantile or progressive epilepsy intellectual disability syndrome* or northern epileps* or progressive myoclonic epileps* or classic juvenile NCL).tw,kf. or ((ceroid storage or CLN or CLN1 or CLN2 or CLN3 or CLN4 or CLN4a or CLN4as or CLN4B or CLN5 or CLN6 or vLINCL or CLN7 or CLN8 or CLN10 or CNCL or CLN11 or CLN12 or CLN13 or Haltia Santavuori or batten mayou or Batten Spielmeyer Vogt Sjogren or "spielmeyer sjogren" or "vogt spielmeyer" or kufs or kuf's or kuf) adj disease*).tw,kf. or (batten adj3 disease*).tw,kf. (3647)

8 exp Phenylketonurias/ or (phenylketonuria* or folling* disease* or biopterin deficienc* or PAH deficienc* or PKU or mpku or mPKU or Non PKU or phenylalanine hydroxylase deficienc* or Mild hyperphenylalaninemia* or Mild HPA or mHPA or Tetrahydrobiopterin responsive hyperphenylalaninemia* or bh4 deficienc* or BH4 responsive or tetrahydrobiopterin responsive HPA or tetrahydrobiopterin deficienc*).tw,kf. (9508)

9 Glycogen Storage Disease Type II/ or (Pompe disease* or Glycogen Storage Disease Type II or Glycogen storage disease type 2 or GSD type 2 or GSD type II or acid maltase dificienc* or Acid alpha-glucosidase deficienc* or Acid alpha glucosidase deficienc* or GAA deficienc* or "Alpha-1,4-glucosidase acid deficienc*" or "Alpha 1,4 glucosidase acid deficienc*" or Glycogenosis type 2 or Glycogenosis type II).tw,kf. or ((GSD adj4 acid maltase deficienc*) or (Glycogenosis adj4 acid maltase deficienc*)).tw,kf. (2831)

10 exp Precursor Cell Lymphoblastic Leukemia-Lymphoma/ or (lymphoblastic leukemia* or lymphoblastic leukaemia* or lymphocytic leukemia* or lymphocytic leukaemia* or lymphoblastic lymphoma* or lymphocytic lymphoma* or precursor B cell lymphoblastic or precursor t cell lymphoblastic or t cell leukemia* or t cell leukaemia* or t lymphocytic leukemia* or t lymphocytic

leukaemia* or precursor cell lymphoblastic or precursor lymphoid neoplasm* or B ALL or T ALL or "childhood all" or "pre b all" or "pre t all").tw,kf. (98052)
11 Cystinosis/ or (cystinos#s or "protein defect of cystin transport" or cystine disease* or cystine diathes#s or cystine storage disease*).tw,kf. (1785)
12 Ataxia-Telangiectasia/ or (ataxia telangiectasia or cerebello-oculocutaneous telangiectasia or cerebello oculocutaneous telangiectasia or Louis Bar* Syndrome*).tw,kf. (8562)
13 Friedreich ataxia/ or (friedreich* adj3 (ataxia* or disease*)).tw,kf. or hereditary spinal scleros#s.tw,kf. or frda.tw,kf. (4266)
14 Arthritis, Juvenile/ or ((juvenile or childhood) adj3 (arthritis or arthritides or spondyloarthropath* or ERA or polyarthritis)).tw,kf. or (still disease* or still's disease* or stills disease*).tw,kf. or ((enthesitis related or enthesitis-related or oligoarticular or polyarticular or psoriasis related or psoriasis-related or psoriatic related or psoriatic-related or systemic onset or systemic-onset or unspecified) adj2 JIA).tw,kf. (18813)
15 Amino Acid Metabolism, Inborn Errors/ or Aromatic-L-Amino-Acid Decarboxylases/df, ge or ((Aromatic L amino acid decarboxylase* or aromatic amino acid decarboxylase* or AADC or DOPA decarboxylase or DDC) adj deficienc*).tw,kf. or AADCD.tw,kf. (6974)
16 Adenosine Deaminase/df, ge or (adenosine deaminase severe combined immunodeficienc* or adenosine deaminase deficienc* or ADA deficienc* or ADA1 deficienc* or ADA-related immune deficienc* or ADA related immune deficienc* or "ADA-SCID" or ADA SCID).tw,kf. (3175)
17 (Aicardi Goutieres Syndrome* or "encephalopathy with basal ganglia calcification" or "Encephalopathy with intracranial calcification" or Pseudotoxoplasmosis syndrome* or Cree encephalitis).tw,kf. (723)
18 alpha-Mannosidosis/ or (alpha-mannosidosis or alpha-mannosidase or mannosidosis or mannosidase or alpha-D-mannosidosis or alpha-D-mannosidase).tw,kf. (3867)
19 Biotinidase Deficiency/ or Amidohydrolases/df, ge or ((Biotinidase or BTD or multiple carboxylase) adj deficienc*).tw,kf. (3745)
20 Duchenne muscular dystrophy/ or ((duchenne or pseudohypertrophic or becker*) adj muscular dystroph*).tw,kf. or dystrophinopath*.tw,kf. (14929)
21 exp Hyperlipoproteinemia Type II/ or (familial hypercholesterolemia* or familial hypercholesterolaemia* or hyperlipoproteinemia type II or hyperlipoproteinemia type 2 or HoFH or ldl receptor disorder*).tw,kf. (11284)
22 Giant Axonal Neuropathy/ or Giant Axonal Neuropath*.tw,kf. (255)
23 lymphohistiocytosis, hemophagocytic/ or (hemophagocytic lymphohistiocytos#s or hemophagocytic syndrome* or genetic HLH or primary HLH or reactive hemophagocytic syndrome* or hemophagocytic reticulos#s or histiocytic reticulos#s).tw,kf. (7876)
24 Hemophilia A/ or (hemophilia A or haemophilia A or autosomal hemophilia or autosomal haemophilia or VIII deficienc* or F8 deficienc*).tw,kf. (26017)
25 Hemophilia B/ or (hemophilia B or haemophilia B or hemophilia BS or haemophilia BS or christmas disease* or F9 deficienc* or factor IX deficienc*).tw,kf. (6103)
26 exp Angioedemas, Hereditary/ or (hereditary angioedema* or C1 INH or C1NH or "complement component 1 inhibitor deficienc*" or "esterase inhibitor deficienc*" or "C1 inhibitor deficienc*" or familial angioneurotic edema* or hereditary angioneurotic edema* or familial angioneurotic oedema* or hereditary angioneurotic oedema* or hereditary bradykinine induced angioedema* or hereditary non histamine induced angioedema* or hereditary non histamine induced angioedema* or "regulatory component deficiency complement component C1" or "HAE with C1 inhibitor deficienc*" or "HAE with C1Inh deficienc*" or "HAE with normal C1 inhibitor" or "HAE with normal C1Inh").tw,kf. (4142)
27 Hypophosphatasia/ or (hypophosphatasia* or rathbun disease* or phosphoethanolaminuria* or odontohypophosphatasia* or alkaline phosphatase deficienc*).tw,kf. (1517)
28 Leukodystrophy, Globoid Cell/ or (Krabbe* disease* or globoid cell leukodystroph* or globoid cell leukoencephalopath*).tw,kf. or ((galactosylcerebrosidase or galactosylceramidase or galc or galactocerebrosidase or galactosylceramide beta galactosidase) adj deficienc*).tw,kf. (1538)

29 Neuroaxonal Dystrophies/ or (neuroaxonal dystroph* or seitelberger* disease* or INAD or INAD1 or phospholipase A2 associated neurodegeneration).tw,kf. (1357)
30 Wolman disease/ or (wolman* disease* or wolman* xanthomatosis or familial xanthomatosis or acid lipase deficienc* or acid lipase disease* or LAL deficienc* or LALD or cholesteryl ester hydrolase dificienc* or cholester#l ester storage disease*).tw,kf. (715)
31 Metal Metabolism, Inborn Errors/ or Sulfite Oxidase/ge, df or (molybdenum cofactor deficienc* or "combined deficiency of sulfite oxidase" or combined molybdoflavoprotein enzyme deficienc* or "combined xanthine oxidase and sulfite oxidase" or "deficiency of molybdenum cofactor" or ISOD or sulfocysteinuria or sulfite oxidase deficienc* or MOCOD or molybdenum cofactor deficienc*).tw,kf. (700)
32 Cryopyrin-Associated Periodic Syndromes/ or (Urticaria/ and cold temperature/) or (NLRP3 associated autoinflammatory disease* or cryopyrinopath* or NLRP3 associated AID or chronic infantile neurological cutaneous or infantile onset multisystem inflammatory disease* or neonatal onset multisystem inflammatory disease* or familial cold urticaria or FCAS or FCU or FCAS1 or kfh or Keratitis fugax hereditarian or Keratoendotheliitis fugax hereditarian or neutrophilic urticaria).tw,kf. or ((cryopyrin associated periodic or CINCA or IOMID or NOMID or Prieur Griscelli or Familial cold autoinflammatory or Muckle Wells or uda or NLRP3 associated autoinflammatory) adj syndrome*).tw,kf. (2605)
33 Neuromyelitis optica/ or (neuromyelitis optica* or devic* disease* or devic* syndrome* or opticomyelitis or NMOSD or opticospinal MS or nmo spectrum disorder*).tw,kf. (7345)
34 Niemann-Pick Disease, Type C/ or (Niemann Pick* Disease* adj2 (type c or type d or nova scotia or "without sphingomyelinase" or "with cholesterol")).tw,kf. (1534)
35 Pantothenate kinase-associated neurodegeneration/ or (pantothenate kinase associated neurodegeneration or NBIA1 or PKAN or pigmentary pallidal degeneration or pigmentary pallidal atroph* or neuroaxonal dystroph*).tw,kf. or (hallervorden spatz adj (disease* or syndrome*)).tw,kf. or ((Neurodegeneration or Neuro-degeneration) adj2 brain iron).tw,kf. or (degeneration adj3 globus).tw,kf. (1744)
36 Hyperoxaluria, Primary/ or (primary hyperoxaluria type 1 or primary hyperoxaluria type I or glycolic aciduria or glyceric aciduria or peroxisomal alanine glyoxylate aminotransferase deficienc*).tw,kf. (1303)
37 Rett Syndrome/ or (cerebroatrophic hyperammonemia* or rtt).tw,kf. or ((rett or retts or rett's) adj (syndrome* or disorder*)).tw,kf. (5261)
38 Severe Combined Immunodeficiency/ or (severe combined immunodeficienc or severe combined immunologic deficienc* or severe combined immune deficienc* or bare lymphocyte syndrome* or familial reticuloendothelios#s or omenn* syndrome*).tw,kf. or ("T+B+ severe combined immunodeficienc*" or "T-B+ severe combined immunodeficienc*" or "T-B+ SCID" or "T-B- severe combined immunodeficienc*" or "T-B- SCID").tw,kf. (5108)
39 Tuberous Sclerosis/ or (tuberous scleros#s or tuberose scleros#s or phakomatos#s or adenoma sebaceum or epiloia or cerebral scleros#s).tw,kf. or (bourneville* adj2 (disease* or syndrome*)).tw,kf. (11694)
40 Tyrosinemias/ or Tyrosine Transaminase/df, ge or (tyrosinemia type 1 or tyrosinemia type i or FAH deficienc* or fumarylacetoacetase deficienc* or fumarylacetoacetate hydrolase deficienc* or hepatorenal tyrosinemia).tw,kf. (1156)
41 Von Hippel-Lindau disease/ or ((van hippel-lindau* or lindau* or VHL) adj (disease* or syndrome*)).tw,kf. or familial cerebelloretinal angiomatos#s.tw,kf. (4507)
42 Hepatolenticular degeneration/ or (wilson* disease* or kinnier-wilson diseas* or cerebral pseudoscleros#s or copper storage disease* or westphal strumpell syndrome* or westphal-strumpell syndrome*).tw,kf. or ((hepatolenticular or hepatocerebral or neurohepatic or hepato-lenticular or hepato-cerebral or neuro-hepatic or progressive lenticular) adj degeneration*).tw,kf. (8980)
43 Adrenoleukodystrophy/ or (adrenoleukodystroph* or X-ALD or X-linked ALD or X-CALD or adrenomyeloneuropath* or bronze schilder disease* or melanodermic leukodystroph* or schilder addison

complex or siemerling creutzfeldt disease*).tw,kf. (2860)
44 Tay-Sachs disease/ or (tay sachs or HEXA deficienc* or HEXA disorder* or hexosaminidase A deficienc* or hexoaminidase alpha subunit deficienc* or beta hexosaminidase subunit alpha deficienc* or GM2 gangliosidosis).tw,kf. (2423)
45 Anemia, Sickle Cell/ or double heterozygotes sickling disorder*.tw,kf. or (sickle cell adj (disease* or anemia* or disorder*)).tw,kf. (32954)
46 Hemiplegia/ or Dystonic Disorders/ge or cerebellar Ataxia/ge or (hemiplegia* or dystonia-parkinsonism or DYT12 or "dystonia 12" or CAPOS syndrome* or "cerebellar ataxia-areflexia-pes cavus-optic" or ATP1A3).tw,kf. or ((cerebellar Ataxia/ or Dystonic Disorders/) and exp mutations/) (21015)
47 Insulin-Like Growth Factor I/df or ((insulin-like or insulin like) adj (growth factor type 1 deficienc* or growth factor type i deficienc* or growth factor deficienc*)).tw,kf. or (growth delay-deafness-intellectual disability or growth delay-hearing loss-intellectual disability or IGF-1 deficienc*).tw,kf. (617)
48 Retinal Dystrophies/ or (retinal dystroph* or inherited retinal disorder*).tw,kf. (3972)
49 Rare Diseases/ or Undiagnosed Diseases/ or ((rare or orphan or undiagnosed) adj (disease* or disorder* or syndrome*)).tw,kf. (63516)
50 (substrate reduction or (substrate adj2 therap*)).tw,kf. (977)
51 exp Enzyme Therapy/ or (enzyme substitution or (enzyme adj2 therap*)).tw,kf. (10574)
52 exp Genetic Therapy/ or ((dna or gametic of gene or genetic) adj therap*).tw,kf. (57045)
53 CRISPR-Cas Systems/ or (crispr-cas9 or crispr-cas9 or crispr-cas or crispr-cas).tw,kf. (42757)
54 Chaperone therap*.tw,kf. (289)
55 or/1-54 (595679)
56 (randomized or placebo or randomly).ab. or clinical trials as topic/ or trial.ti. or (randomized controlled trial or controlled clinical trial).pt. (1664322)
57 (adaptive clinical trial or clinical trial or clinical trial, phase i or clinical trial, phase ii or clinical trial, phase iii or clinical trial, phase iv or controlled clinical trial or clinical trial protocol or pragmatic clinical trial or randomized controlled trial).pt. (1031622)
58 clinical trials as topic/ or adaptive clinical trials as topic/ or Clinical Studies as Topic/ or clinical trials as topic/ or clinical trials, phase i as topic/ or clinical trials, phase ii as topic/ or clinical trials, phase iii as topic/ or clinical trials, phase iv as topic/ or controlled clinical trials as topic/ or non-randomized controlled trials as topic/ or randomized controlled trials as topic/ or double-blind method/ or pragmatic clinical trials as topic/ or intention to treat analysis/ or early termination of clinical trials/ or Controlled Before-After Studies/ or cross-over studies/ or Clinical Trial Protocol/ (601065)
59 (((Before-After or "before and after" or pre-post or pre post) adj (studies or study or trial or trials)) or master protocol trial*).tw,kf. (7644)
60 ((randomi#ation or randomi#ed or non-randomi#ed or controlled or "delayed start" or self-controlled or single-arm or open label or clinical or (cross adj over) or crossover or pragmatic or factorial or basket or platform or "n-of-1" or "n of 1" or "n-of-few" or "single participant" or "single patient") adj3 (studies or study or trial or trials)).tw,kf. (1531521)
61 or/56-60 (2617930)
62 55 and 61 (53457)
63 (animals/ not humans/) or comment/ or editorial/ or hi.fs. or case report.mp. (7523110)
64 62 not 63 (50571)
65 limit 64 to yr="2014 -Current" (25515)
66 exp Infant/ or pediatrics/ or Adolescent/ or exp child/ or (Infan* or newborn* or new-born* or perinat* or neonat* or baby* or babies or toddler* or minors* or boy or boys or boyfriend or boyhood or girl* or kid or kids or child* or schoolchild* or adolescen* or juvenil* or youth* or teen* or under*age* or pubescen* or pediatric* or paediatric* or school* or prematur* or preterm*).tw,kf. (5388771)
67 65 and 66 (9557)
68 limit 65 to "all child (0 to 18 years)" (6966)
69 67 or 68 (9557)

Database: EBM Reviews - Cochrane Central Register of Controlled Trials <January 2025>

---

1 (((fraxa or "fra(x)" or frax or fraxe or FXS or marker x or martin bell) adj syndrome*) or FRAXA or (fragile X adj3 syndrome*)).tw. (198)

2 (fabry* disease* or alpha galactosidase a deficienc* or anderson fabry diseas* or angiokeratoma* or GLA deficienc* or hereditary dystopic lipidos* or ceramide trihexosidase deficienc*).tw. (257)

3 (cystic fibrosis or fibrocystic disease* or mucoviscidosis or pancreas* fibrosis or pancrea* fibrocystic disease*).tw. (6046)

4 (spinal muscular atroph* or spinal amyotroph* or SMA type 1 or SMA type I or SMA-I or SMA1 or SMA type 2 or SMA type II or SMA-II or SMA2 or SMA type 3 or SMA type III or SMA-III or SMA3 or "SMN related SMA" or chromosome 5 SMA or neurogenic scapuloperonealamyotroph* or progressive muscular atroph* or myelopathic atroph* or bulbospinal neuronopath* or hereditary motor neuronopath* or myelopathic muscular atroph* or ((dubowitz or Werdnig Hoffmann or kugelberg welander) adj disease*)).tw. (339)

5 (mucopolysaccharidos* or MPS disorder* or lipochondrodystroph* or polydystrophic dwarfism or ((alpha-L-iduronidase or Iduronate 2-sulfatase or i2s or Hyaluronidase or Beta-glucuronidase or ARSB or ASB or Arylsulfatase B or N-acetylgalactosamine 4-sulfatase or gusb) adj deficienc*) or ((hurler* or pfaundler-hurler or scheie or hurler-scheie or hunter* or Sanfilippo or Morquio or Sly or Maroteaux-Lamy or Maroteaux) adj2 (diseas* or syndrome*))).tw. (215)

6 (non-neuronopathic GD or neuronopathic GD or Glucocerebrosidase deficienc* or Glucosylceramidase deficienc* or Acid beta glucosidase deficienc* or glucosyl cerebroside lipidos#s or glucosylceramide beta glucosidase deficienc* or glucosylceramide lipidos#s or kerasinhistiocytos#s or lipoidos#s histiocytos* or kerasin thesaurismos#s or ((gaucher* or Gaucher-like) adj (diseas* or syndrome*))).tw. (209)

7 (neuronal ceroid-lipofuscinos* or neuronal ceroid lipofuscinos* or cerebroretinal degeneration or parkinsonism-neuronal ceroid lipofuscinos* or parkinsonism neuronal ceroid lipofuscinos* or turkish variant late infantile or progressive epilepsy intellectual disability syndrome* or northern epileps* or progressive myoclonic epileps* or classic juvenile NCL or ((ceroid storage or CLN or CLN1 or CLN2 or CLN3 or CLN4 or CLN4a or CLN4as or CLN4B or CLN5 or CLN6 or vLINCL or CLN7 or CLN8 or CLN10 or CNCL or CLN11 or CLN12 or CLN13 or Haltia Santavuori or batten mayou or Batten Spielmeyer Vogt Sjogren or "spielmeyer sjogren" or "vogt spielmeyer" or kufs or kuf's or kuf) adj disease*) or (batten adj3 disease*)).tw. (31)

8 (phenylketonuria* or folling* disease* or biopterin deficienc* or PAH deficienc* or PKU or mpku or mPKU or Non PKU or phenylalanine hydroxylase deficienc* or Mild hyperphenylalaninemia* or Mild HPA or mHPA or Tetrahydrobiopterin responsive hyperphenylalaninemia* or bh4 deficienc* or BH4 responsive or tetrahydrobiopterin responsive HPA or tetrahydrobiopterin deficienc*).tw. (448)

9 (Pompe disease* or Glycogen Storage Disease Type II or Glycogen storage disease type 2 or GSD type 2 or GSD type II or acid maltase dificienc* or Acid alpha-glucosidase deficienc* or Acid alpha glucosidase deficienc* or GAA deficienc* or "Alpha-1,4-glucosidase acid deficienc*" or "Alpha 1,4 glucosidase acid deficienc*" or Glycogenosis type 2 or Glycogenosis type II or ((GSD adj4 acid maltase deficienc*) or (Glycogenosis adj4 acid maltase deficienc*))).tw. (150)

10 (lymphoblastic leukemia* or lymphoblastic leukaemia* or lymphocytic leukemia* or lymphocytic leukaemia* or lymphoblastic lymphoma* or lymphocytic lymphoma* or precursor B cell lymphoblastic or precursor t cell lymphoblastic or t cell leukemia* or t cell leukaemia* or t lymphocytic leukemia* or t lymphocytic leukaemia* or precursor cell lymphoblastic or precursor lymphoid neoplasm* or B ALL or T ALL or "childhood all" or "pre b all" or "pre t all").tw. (6620)

11 (cystinos#s or "protein defect of cystin transport" or cystine disease* or cystine diathes#s or cystine storage disease*).tw. (46)

12 (ataxia telangiectasia or cerebello-oculocutaneous telangiectasia or cerebello oculocutaneous

telangiectasia or Louis Bar* Syndrome*).tw. (73)
13 ((friedreich* adj3 (ataxia* or disease*)) or hereditary spinal scleros#s or frda).tw. (173)
14 (((juvenile or childhood) adj3 (arthritis or arthritides or spondyloarthropath* or ERA or polyarthritis)) or (still disease* or still's disease* or stills disease*) or ((enthesitis related or enthesitis-related or oligoarticular or polyarticular or psoriasis related or psoriasis-related or psoriatic related or psoriatic-related or systemic onset or systemic-onset or unspecified) adj2 JIA)).tw. (1069)
15 (((Aromatic L amino acid decarboxylase* or aromatic amino acid decarboxylase* or AADC or DOPA decarboxylase or DDC) adj deficienc*) or AADCD).tw. (3)
16 (adenosine deaminase severe combined immunodeficienc* or adenosine deaminase deficienc* or ADA deficienc* or ADA1 deficienc* or ADA-related immune deficienc* or ADA related immune deficienc* or "ADA-SCID" or ADA SCID).tw. (4)
17 (Aicardi Goutieres Syndrome* or "encephalopathy with basal ganglia calcification" or "Encephalopathy with intracranial calcification" or Pseudotoxoplasmosis syndrome* or Cree encephalitis).tw. (1)
18 (alpha-mannosidosis or alpha-mannosidase or mannosidosis or mannosidase or alpha-D-mannosidosis or alpha-D-mannosidase).tw. (21)
19 ((Biotinidase or BTD or multiple carboxylase) adj deficienc*).tw. (4)
20 (((duchenne or pseudohypertrophic or becker*) adj muscular dystroph*) or dystrophinopath*).tw. (895)
21 (familial hypercholesterolemia* or familial hypercholesterolaemia* or hyperlipoproteinemia type II or hyperlipoproteinemia type 2 or HoFH or ldl receptor disorder*).tw. (997)
22 Giant Axonal Neuropath*.tw. (1)
23 (hemophagocytic lymphohistiocytos#s or hemophagocytic syndrome* or genetic HLH or primary HLH or reactive hemophagocytic syndrome* or hemophagocytic reticulos#s or histiocytic reticulos#s).tw. (63)
24 (hemophilia A or haemophilia A or autosomal hemophilia or autosomal haemophilia or VIII deficienc* or F8 deficienc*).tw. (1031)
25 (hemophilia B or haemophilia B or hemophilia BS or haemophilia BS or christmas disease* or F9 deficienc* or factor IX deficienc*).tw. (244)
26 (hereditary angioedema* or C1 INH or C1NH or "complement component 1 inhibitor deficienc*" or "esterase inhibitor deficienc*" or "C1 inhibitor deficienc*" or familial angioneurotic edema* or hereditary angioneurotic edema* or familial angioneurotic oedema* or hereditary angioneurotic oedema* or hereditary bradykinine induced angioedema* or hereditary non histamine induced angioedema* or hereditary non histamine induced angioedema* or "regulatory component deficiency complement component C1" or "HAE with C1 inhibitor deficienc*" or "HAE with C1Inh deficienc*" or "HAE with normal C1 inhibitor" or "HAE with normal C1Inh").tw. (523)
27 (hypophosphatasia* or rathbun disease* or phosphoethanolaminuria* or odontohypophosphatasia* or alkaline phosphatase deficienc*).tw. (42)
28 (Krabbe* disease* or globoid cell leukodystroph* or globoid cell leukoencephalopath* or ((galactosylcerebrosidase or galactosylceramidase or galc or galactocerebrosidase or galactosylceramide b eta galactosidase) adj deficienc*)).tw. (1)
29 (neuroaxonal dystroph* or seitelberger* disease* or INAD or INAD1 or phospholipase A2 associated neurodegeneration).tw. (9)
30 (wolman* disease* or wolman* xanthomatosis or familial xanthomatosis or acid lipase deficienc* or acid lipase disease* or LAL deficienc* or LALD or cholesteryl ester hydrolase dificienc* or cholester#l ester storage disease*).tw. (35)
31 (molybdenum cofactor deficienc* or "combined deficiency of sulfite oxidase" or combined molybdoflavoprotein enzyme deficienc* or "combined xanthine oxidase and sulfite oxidase" or "deficiency of molybdenum cofactor" or ISOD or sulfocysteinuria or sulfite oxidase deficienc* or MOCOD or molybdenum cofactor deficienc*).tw. (6)
32 (NLRP3 associated autoinflammatory disease* or cryopyrinopath* or NLRP3 associated AID or

chronic infantile neurological cutaneous or infantile onset multisystem inflammatory disease* or neonatal onset multisystem inflammatory disease* or familial cold urticaria or FCAS or FCU or FCAS1 or kfh or Keratitis fugax hereditarian or Keratoendotheliitis fugax hereditarian or neutrophilic urticaria or ((cryopyrin associated periodic or CINCA or IOMID or NOMID or Prieur Griscelli or Familial cold autoinflammatory or Muckle Wells or uda or NLRP3 associated autoinflammatory) adj syndrome*)).tw. (179)
33 (neuromyelitis optica* or devic* disease* or devic* syndrome* or opticomyelitis or NMOSD or opticospinal MS or nmo spectrum disorder*).tw. (394)
34 (Niemann Pick* Disease* adj2 (type c or type d or nova scotia or "without sphingomyelinase" or "with cholesterol")).tw. (36)
35 (pantothenate kinase associated neurodegeneration or NBIA1 or PKAN or pigmentary pallidal degeneration or pigmentary pallidal atroph* or neuroaxonal dystroph* or (hallervorden spatz adj (disease* or syndrome*)) or ((Neurodegeneration or Neuro-degeneration) adj2 brain iron) or (degeneration adj3 globus)).tw. (26)
36 (primary hyperoxaluria type 1 or primary hyperoxaluria type I or glycolic aciduria or glyceric aciduria or peroxisomal alanine glyoxylate aminotransferase deficienc*).tw. (44)
37 (cerebroatrophic hyperammonemia* or rtt or ((rett or retts or rett's) adj (syndrome* or disorder*))).tw. (191)
38 (severe combined immunodeficienc or severe combined immunologic deficienc* or severe combined immune deficienc* or bare lymphocyte syndrome* or familial reticuloendothelios#s or omenn* syndrome* or ("T+B+ severe combined immunodeficienc*" or "T-B+ severe combined immunodeficienc*" or "T-B+ SCID" or "T-B- severe combined immunodeficienc*" or "T-B- SCID")).tw. (1)
39 (tuberous scleros#s or tuberose scleros#s or phakomatos#s or adenoma sebaceum or epiloia or cerebral scleros#s or (bourneville* adj2 (disease* or syndrome*))).tw. (294)
40 (tyrosinemia type 1 or tyrosinemia type i or FAH deficienc* or fumarylacetoacetase deficienc* or fumarylacetoacetate hydrolase deficienc* or hepatorenal tyrosinemia).tw. (6)
41 (((van hippel-lindau* or lindau* or VHL) adj (disease* or syndrome*)) or familial cerebelloretinal angiomatos#s).tw. (22)
42 (wilson* disease* or kinnier-wilson diseas* or cerebral pseudoscleros#s or copper storage disease* or westphal strumpell syndrome* or westphal-strumpell syndrome* or ((hepatolenticular or hepatocerebral or neurohepatic or hepato-lenticular or hepato-cerebral or neuro-hepatic or progressive lenticular) adj degeneration*)).tw. (193)
43 (adrenoleukodystroph* or X-ALD or X-linked ALD or X-CALD or adrenomyeloneuropath* or bronze schilder disease* or melanodermic leukodystroph* or schilder addison complex or siemerling creutzfeldt disease*).tw. (53)
44 (tay sachs or HEXA deficienc* or HEXA disorder* or hexosaminidase A deficienc* or hexoaminidase alpha subunit deficienc* or beta hexosaminidase subunit alpha deficienc* or GM2 gangliosidosis).tw. (15)
45 (double heterozygotes sickling disorder* or (sickle cell adj (disease* or anemia* or disorder*))).tw. (2132)
46 (hemiplegia* or dystonia-parkinsonism or DYT12 or "dystonia 12" or CAPOS syndrome* or "cerebellar ataxia-areflexia-pes cavus-optic" or ATP1A3).tw. (2164)
47 (((insulin-like or insulin like) adj (growth factor type 1 deficienc* or growth factor type i deficienc* or growth factor deficienc*)) or (growth delay-deafness-intellectual disability or growth delay-hearing loss-intellectual disability or IGF-1 deficienc*)).tw. (8)
48 (retinal dystroph* or inherited retinal disorder*).tw. (40)
49 ((rare or orphan or undiagnosed) adj (disease* or disorder* or syndrome*)).tw. (961)
50 (substrate reduction or (substrate adj2 therap*)).tw. (58)
51 (enzyme substitution or (enzyme adj2 therap*)).tw. (855)
52 ((dna or gametic of gene or genetic) adj therap*).tw. (15)

53 (crispr-cas9 or crispr-cas9 or crispr-cas or crispr-cas).tw. (24)
54 Chaperone therap*.tw. (10)
55 or/1-54 (26514)
56 (Infan* or newborn* or new-born* or perinat* or neonat* or baby* or babies or toddler* or minors* or boy or boys or boyfriend or boyhood or girl* or kid or kids or child* or schoolchild* or adolescen* or juvenil* or youth* or teen* or under*age* or pubescen* or pediatric* or paediatric* or school* or prematur* or preterm*).tw. (291180)
57 55 and 56 (9292)
58 limit 57 to yr="2014 -Current" (5112)

eMethods 2. Data items extracted from all articles that met the initial inclusion criteria (N=67)

Trial registration

Trial publication year

Document type

Study type

Country of internal arm sites

Funding

Trial phase

Extension component

Internal arm(s)

- Number
- Allocation
- Sample size
- Minimum/Maximum age eligibility
- Sex and/or gender eligibility
- Equity, diversity and inclusion language

Rare disease(s)

eMethods 3. Data items extracted from the subset of articles that met the additional inclusion criteria for detailed extraction (n=18)

Protocol availability

Number of internal arms compared to an external arm

Internal arm(s) – collected for each internal arm

- Type of arm
- Intervention
- Recruitment period
- Data collection period
- Total length of follow up
- Inclusion criteria
- Sex at birth included
- Minimum/Maximum age included
- Sample size

External control arm(s) – collected for each external arm

- Pre-planification
- Rationale provided
- Number of arms
- Data source type
- Recruitment period
- Data collection period
- Total length of follow up
- Accordance of inclusion criteria with that of internal arm
- Sex at birth included
- Minimum/Maximum age included
- Sample size
- Countries of study sites

Analysis comparing the internal and external arms

- Analytic method used to address covariate imbalance
- Comparison of baseline characteristics
- Number of outcomes compared using individual-level data

- Outcomes compared using individual-level data – collected for each outcome
    - Description (instrument/definition)
    - Metric
    - Primary or secondary/exploratory
    - Internal and external arms included in comparison
    - Assessment time points used in analysis for internal arm(s)
    - Assessment time points used in analysis for external arm(s)
    - Blinding in internal arm(s)
    - Blinding in external arm(s)
    - Statistical method

Reporting

- Internal arm(s)
    - Loss to follow-up in internal arm(s)
    - Modifications or non-adherence to treatment
    - Missingness in baseline characteristics
    - Missingness in outcome measurements
- External arm(s)
    - Missingness in baseline characteristics
    - Missingness in outcome measurements
    - Pre-specification of how missingness would be handled
- Sensitivity analysis including external control data
- Quantitative bias analyses specific to the external comparison

eTable 1. Articles included in the scoping review

First author and year of publication	Article title	External control data type	Attempted to address baseline differences[a]
Clemens 2020[57]	Safety, tolerability, and efficacy of viltolarsen in boys with Duchenne muscular dystrophy amenable to exon 53 skipping: A phase 2 randomized clinical trial	Individual	Yes
Clemens 2022[61]	Long-term functional efficacy and safety of viltolarsen in patients with Duchenne muscular dystrophy	Individual	Yes
Clemens 2023[64]	Efficacy and safety of viltolarsen in boys with Duchenne muscular dystrophy: Results from the phase 2, open-label, 4-year extension study	Individual	Yes
Day 2021[59]	Onasemnogene abeparvovec gene therapy for symptomatic infantile-onset spinal muscular atrophy in patients with two copies of SMN2 (STR1VE): An open-label, single-arm, multicentre, phase 3 trial	Individual	Yes
Fumagalli 2022[62]	Lentiviral haematopoietic stem-cell gene therapy for early-onset metachromatic leukodystrophy: Long-term results from a non-randomised, open-label, phase 1/2 trial and expanded access	Individual	Yes
Harper 2024[68]	Safety and efficacy of viltolarsen in ambulatory and nonambulatory males with Duchenne muscular dystrophy	Individual	Yes
Hendriksz 2016[52]	Long-term endurance and safety of elosulfase alfa enzyme replacement therapy in patients with Morquio A syndrome	Individual	Yes
Khan 2019[56]	Eteplirsen treatment attenuates respiratory decline in ambulatory and non-ambulatory patients with Duchenne muscular dystrophy	Individual	Yes
Konstan 2017[53]	Assessment of safety and efficacy of long-term treatment with combination lumacaftor and ivacaftor therapy in patients with cystic fibrosis homozygous for the F508del-CFTR mutation (PROGRESS): A phase 3, extension study	Individual	Yes
Krosschell 2018[54]	Clinical trial of L-carnitine and valproic acid in spinal muscular atrophy type I	Individual	Yes
Mah 2022[63]	Efficacy and safety of vamorolone in Duchenne muscular dystrophy: A 30-month nonrandomized controlled open-label extension trial	Individual	Yes
McDonald 2021[60]	Open-label evaluation of eteplirsen in patients with Duchenne muscular dystrophy amenable to exon 51 skipping: PROMOVI trial	Individual	Yes
Mercuri 2023[65]	Risdiplam in types 2 and 3 spinal muscular atrophy: A randomised, placebo-controlled, dose-finding trial followed by 24 months of treatment	Individual	Yes

Muntoni 2020[58]	Long-term follow-up of patients with type 2 and non-ambulant type 3 spinal muscular atrophy (SMA) treated with olesoxime in the OLEOS trial	Individual	Yes
Pignolo 2023[66]	Reduction of new heterotopic ossification (HO) in the open-label, phase 3 MOVE trial of palovarotene for fibrodysplasia ossificans progressiva (FOP)	Individual	Yes
Schulz 2018[55]	Study of intraventricular cerliponase alfa for CLN2 disease	Individual	Yes
Schulz 2024[69]	Safety and efficacy of cerliponase alfa in children with neuronal ceroid lipofuscinosis type 2 (CLN2 disease): An open-label extension study	Individual	Yes
Zaidman 2023[67]	Delandistrogene moxeparvovec gene therapy in ambulatory patients (aged ≥4 to <8 Years) with Duchenne muscular dystrophy: 1-year interim results from study SRP-9001-103 (ENDEAVOR)	Individual	Yes
Chiriboga 2024[101]	JEWELFISH: 24-month results from an open-label study in non-treatment-naive patients with SMA receiving treatment with risdiplam	Individual	No
Crawford 2023[97]	Continued benefit of nusinersen initiated in the presymptomatic stage of spinal muscular atrophy: 5-year update of the NURTURE study	Individual	No
Connolly 2019[80]	Twice-weekly glucocorticosteroids in infants and young boys with Duchenne muscular dystrophy	Individual	No
Darras 2021[87]	Risdiplam-treated infants with type 1 spinal muscular atrophy versus historical controls	Individual	No
Deiva 2021[88]	Intracerebral gene therapy in four children with Sanfilippo B syndrome: 5.5-year follow-up results	Individual	No
DeVivo 2019[81]	Nusinersen initiated in infants during the presymptomatic stage of spinal muscular atrophy: Interim efficacy and safety results from the Phase 2 NURTURE study	Individual	No
Finkel 2016[73]	Treatment of infantile-onset spinal muscular atrophy with nusinersen: A phase 2, open-label, dose-escalation study	Individual	No
Gentner 2021[89]	Hematopoietic stem- and progenitor-cell gene therapy for Hurler syndrome	Individual	No
Ghosh 2021[90]	High dose genistein in Sanfilippo syndrome: A randomised controlled trial	Individual	No
Hacein-Bey-Abina 2014[70]	A modified gamma-retrovirus vector for X-linked severe combined immunodeficiency	Individual	No
Henzi 2023[98]	Safety and efficacy of tamoxifen in boys with Duchenne muscular dystrophy (TAMDMD): A multicentre, randomised, double-blind, placebo-controlled, phase 3 trial	Individual	No
Hoffman 2019[82]	Vamorolone trial in Duchenne muscular dystrophy shows dose-related improvement of muscle function	Individual	No
Jones 2015[72]	Safety and clinical activity of elosulfase alfa in pediatric patients with Morquio A syndrome (mucopolysaccharidosis IVA) less than 5 y	Individual	No
Jones 2017[76]	Survival in infants treated with sebelipase Alfa for lysosomal acid lipase deficiency: An open-label, multicenter, dose-escalation study	Individual	No
Kato 2022[93]	Sirolimus for epileptic seizures associated with focal cortical dysplasia type II	Individual	No

Kinane 2018[79]	Long-term pulmonary function in Duchenne muscular dystrophy: Comparison of eteplirsen-treated patients to natural history	Individual	No
Levin 2014[71]	Oral cysteamine bitartrate and N-acetylcysteine for patients with infantile neuronal ceroid lipofuscinosis: A pilot study	Individual	No
Magnani 2022[94]	Long-term safety and efficacy of lentiviral hematopoietic stem/progenitor cell gene therapy for Wiskott-Aldrich syndrome	Individual	No
Ory 2017[77]	Intrathecal 2-hydroxypropyl-beta-cyclodextrin decreases neurological disease progression in Niemann-Pick disease, type C1: A non-randomised, open-label, phase 1-2 trial	Individual	No
Polgreen 2020[83]	Clinical trial of laronidase in Hurler syndrome after hematopoietic cell transplantation	Individual	No
Seo 2023[99]	Intracerebroventricular enzyme replacement therapy in patients with neuronopathic mucopolysaccharidosis type II: Final report of 5-year results from a Japanese open-label phase 1/2 study	Individual	No
Shieh 2023[100]	Safety and efficacy of gene replacement therapy for X-linked myotubular myopathy (ASPIRO): A multinational, open-label, dose-escalation trial	Individual	No
Smith 2020[84]	Efficacy and safety of vamorolone in Duchenne muscular dystrophy: An 18-month interim analysis of a non-randomized open-label extension study	Individual	No
Sondhi 2020[85]	Slowing late infantile Batten disease by direct brain parenchymal administration of a rh.10 adeno-associated virus expressing CLN2	Individual	No
Sondhi 2024[102]	Twenty-year survival analysis of adeno-associated virus vector serotype 2-mediated gene therapy to the central nervous system for CLN2 disease	Individual	No
Strauss 2022[95]	Onasemnogene abeparvovec for presymptomatic infants with two copies of SMN2 at risk for spinal muscular atrophy type 1: The phase III SPR1NT trial	Individual	No
Strauss 2022[96]	Onasemnogene abeparvovec for presymptomatic infants with three copies of SMN2 at risk for spinal muscular atrophy: The phase III SPR1NT trial	Individual	No
Tardieu 2017[78]	Intracerebral gene therapy in children with mucopolysaccharidosis type IIIB syndrome: An uncontrolled phase 1/2 clinical trial	Individual	No
Tsabari 2021[91]	Safety and clinical outcome of tamoxifen in Duchenne muscular dystrophy	Individual	No
Vijay 2021[92]	Long-term survival with sebelipase alfa enzyme replacement therapy in infants with rapidly progressive lysosomal acid lipase deficiency: Final results from 2 open-label studies	Individual	No
Wagner 2020[86]	Randomized phase 2 trial and open-label extension of domagrozumab in Duchenne muscular dystrophy	Individual	No
Whyte 2016[74]	Asfotase alfa therapy for children with hypophosphatasia	Individual	No
Whyte 2016[75]	Asfotase alfa treatment improves survival for perinatal and infantile hypophosphatasia	Individual	No

Auerswald 2015[36]	The EPIC study: A lesson to learn	Aggregate	NA
Chien 2017[37]	Efficacy and safety of AAV2 gene therapy in children with aromatic L-amino acid decarboxylase deficiency: An open-label, phase 1/2 trial	Aggregate	NA
Dvorak 2014[38]	A trial of alemtuzumab adjunctive therapy in allogeneic hematopoietic cell transplantation with minimal conditioning for severe combined immunodeficiency	Aggregate	NA
Gilchrist 2024[39]	ALPINE2: Efficacy and safety of 14-day vs 28-day inhaled aztreonam for Pa eradication in children with cystic fibrosis	Aggregate	NA
Hankins 2014[40]	From infancy to adolescence: Fifteen years of continuous treatment with hydroxyurea in sickle cell anemia	Aggregate	NA
Hoppe 2024[41]	Vanzacaftor-tezacaftor-deutivacaftor for children aged 6-11 years with cystic fibrosis (RIDGELINE Trial VX21-121-105): An analysis from a single-arm, phase 3 trial	Aggregate	NA
Kavakli 2015[42]	Prophylaxis vs. on-demand treatment with BAY 81-8973, a full-length plasma protein-free recombinant factor VIII product: Results from a randomized trial (LEOPOLD II)	Aggregate	NA
Lynch 2019[44]	Randomized, double-blind, placebo-controlled study of interferon-γ 1b in Friedreich Ataxia	Aggregate	NA
Lynch 2019[43]	Safety, pharmacodynamics, and potential benefit of omaveloxolone in Friedreich ataxia	Aggregate	NA
Manco-Johnson 2016[45]	Efficacy and safety of protein C concentrate to treat purpura fulminans and thromboembolic events in severe congenital protein C deficiency	Aggregate	NA
Mendell 2017[46]	Single-dose gene-replacement therapy for spinal muscular atrophy	Aggregate	NA
Mendell 2020[47]	Assessment of systemic delivery of rAAVrh74.MHCK7.micro-dystrophin in children with Duchenne muscular dystrophy: A nonrandomized controlled trial	Aggregate	NA
Patterson 2024[48]	Disease-modifying, neuroprotective effect of N-acetyl-L-leucine in adult and pediatric patients with Niemann-Pick disease type C	Aggregate	NA
Saxena 2016[49]	Efficacy and safety of BAY 81-8973, a full-length recombinant factor VIII: Results from the LEOPOLD I trial	Aggregate	NA
Taylor 2019[50]	Cardiac and skeletal muscle effects in the randomized HOPE-Duchenne trial	Aggregate	NA
Windyga 2014[51]	Pharmacokinetics, efficacy and safety of BAX326, a novel recombinant factor IX: A prospective, controlled, multicentre phase I/III study in previously treated patients with severe (FIX level <1%) or moderately severe (FIX level ≤2%) haemophilia B	Aggregate	NA

[a]Beyond age and baseline score of outcome.

NA: Not applicable

eTable 2. Genetic RDs represented across all included articles

Rare disease	Overall (N=67)	Other (n=49)	Prioritized subset[a] (n=18)
Duchenne muscular dystrophy	17 (21%)	9 (15%)	8 (36%)
Spinal muscular atrophy			
Spinal muscular atrophy	2 (2.5%)	2 (3.4%)	0 (0%)
Spinal muscular atrophy type 1	9 (11%)	6 (10%)	3 (14%)
Spinal muscular atrophy type 2	6 (7.4%)	4 (6.8%)	2 (9.1%)
Spinal muscular atrophy type 3	3 (3.7%)	1 (1.7%)	2 (9.1%)
Cystic Fibrosis	3 (3.7%)	2 (3.4%)	1 (4.5%)
Hemophilia			
Severe hemophilia A	3 (3.7%)	3 (5.1%)	0 (0%)
Severe hemophilia B	1 (1.2%)	1 (1.7%)	0 (0%)
Moderate hemophilia A	1 (1.2%)	1 (1.7%)	0 (0%)
Moderate hemophilia B	1 (1.2%)	1 (1.7%)	0 (0%)
Mucopolysaccharidosis			
Mucopolysaccharidosis type 1H	2 (2.5%)	2 (3.4%)	0 (0%)
Mucopolysaccharidosis type 2	1 (1.2%)	1 (1.7%)	0 (0%)
Mucopolysaccharidosis type 3A	1 (1.2%)	1 (1.7%)	0 (0%)
Mucopolysaccharidosis type 3B	3 (3.7%)	3 (5.1%)	0 (0%)
Mucopolysaccharidosis type 3C	1 (1.2%)	1 (1.7%)	0 (0%)
Mucopolysaccharidosis type 4A	2 (2.5%)	1 (1.7%)	1 (4.5%)
Ceroid lipofuscinosis			
Neuronal ceroid lipofuscinosis type 2 (CLN2)	2 (2.5%)	0 (0%)	2 (9.1%)
Late infantile neuronal ceroid lipofuscinosis type 2	2 (2.5%)	2 (3.4%)	0 (0%)
Infantile neuronal ceroid lipofuscinosis type 1	1 (1.2%)	1 (1.7%)	0 (0%)
Friedreich ataxia	2 (2.5%)	2 (3.4%)	0 (0%)
Lysosomal acid lipase deficiency	2 (2.5%)	2 (3.4%)	0 (0%)
Niemann-Pick disease type C	2 (2.5%)	2 (3.4%)	0 (0%)
Aromatic L-amino acid decarboxylase deficiency	1 (1.2%)	1 (1.7%)	0 (0%)
Fibrodysplasia ossificans progressiva	1 (1.2%)	0 (0%)	1 (4.5%)
Hypophosphatasia			
Hypophosphatasia	1 (1.2%)	1 (1.7%)	0 (0%)
Childhood-onset hypophosphatasia	1 (1.2%)	1 (1.7%)	0 (0%)
Infantile hypophosphatasia	1 (1.2%)	1 (1.7%)	0 (0%)
Isolated focal cortical dysplasia type 2	1 (1.2%)	1 (1.7%)	0 (0%)

Rare disease	Overall (N=67)	Other (n=49)	Prioritized subset[a] (n=18)
Metachromatic leukodystrophy			
Juvenile metachromatic leukodystrophy	1 (1.2%)	0 (0%)	1 (4.5%)
Late infantile metachromatic leukodystrophy	1 (1.2%)	0 (0%)	1 (4.5%)
Severe combined immunodeficiency			
Severe combined immunodeficiency	1 (1.2%)	1 (1.7%)	0 (0%)
X-linked severe combined immunodeficiency	1 (1.2%)	1 (1.7%)	0 (0%)
Severe hereditary thrombophilia due to congenital protein C deficiency	1 (1.2%)	1 (1.7%)	0 (0%)
Sickle cell anemia	1 (1.2%)	1 (1.7%)	0 (0%)
Wiskott-Aldrich syndrome	1 (1.2%)	1 (1.7%)	0 (0%)
X-linked centronuclear myopathy	1 (1.2%)	1 (1.7%)	0 (0%)

Note. The count of genetic RD/RD subtypes is greater than 67 as several articles included individuals with more than one subtype of RD.

[a]Prioritized subset: used individual-level data for external control group(s) and reported using an analytic method to address baseline covariate imbalances beyond adjusting for age and baseline score of outcome.

eTable 3. Related externally controlled trials that attempted to address differences in baseline covariates between arms.

Covidence ID (citation)	Article title	Linked articles	Intervention	External control
Cystic fibrosis				
#4461 (Konstan et al., 2017)[53]	Assessment of safety and efficacy of long-term treatment with combination lumacaftor and ivacaftor therapy in patients with cystic fibrosis homozygous for the F508del-CFTR mutation (PROGRESS): A phase 3, extension study	None	Lumacaftor and Ivacaftor	No other articles with the same source of external data.
Duchenne muscular dystrophy				
#101 (Clemens et al., 2020)[57]	Safety, tolerability, and efficacy of viltolarsen in boys with Duchenne muscular dystrophy amenable to exon 53 skiping: A phase 2 randomized clinical trial	#2497 (interim analysis of extension) and #47 (final report of extension).	Viltolarsen	All 8 articles included data from the Cooperative International Neuromuscular Research Group Duchenne Natural History Study in their external control
#2497 (Clemens et al., 2022)[61]	Long-term functional efficacy and safety of viltolarsen in patients with Duchenne muscular dystrophy	#101 (primary report) and #47	Viltolarsen	
#47 (Clemens et al., 2023)[64]	Efficacy and safety of viltolarsen in boys with Duchenne muscular dystrophy: Results from the phase 2, open-label, 4-year extension study	#101 and #2497	Viltolarsen	
#6120 (Harper et al., 2024)[68]	Safety and efficacy of viltolarsen in ambulatory and nonambulatory males with Duchenne muscular dystrophy	None	Viltolarsen	
#5727 (Khan et al., 2019)[56]	Eteplirsen treatment attenuates respiratory decline in ambulatory and non-ambulatory patients with Duchenne muscular dystrophy	#3701 (full analysis of one of the internal arms [study 301])	Eteplirsen	
#8130 (McDonald et al., 2021)[60]	Open-label evaluation of eteplirsen in patients with Duchenne muscular dystrophy amenable to exon 53 skipping: PROMOVI trial	#5727 (one internal arm included interim data from study 301)	Eteplirsen	
#3701 (Mah et al., 2022)[63]	Efficacy and safety of vamorolone in Duchenne muscular dystrophy: A 30-month nonrandomized controlled open-label extension trial	None	Vamorolone	
#15928 (Zaidman et al., 2023)[67]	Delandistrogene moxeparvovec gene therapy in ambultory patients (aged ≥4 to <8 years) in Duchenne musuclar dystrophy: 1-year interim results from study SRP-9001-103 (ENDEAVOR)	None	Delandistrogene moxeparvovec	
Fibrodysplasia ossificans progressiva				

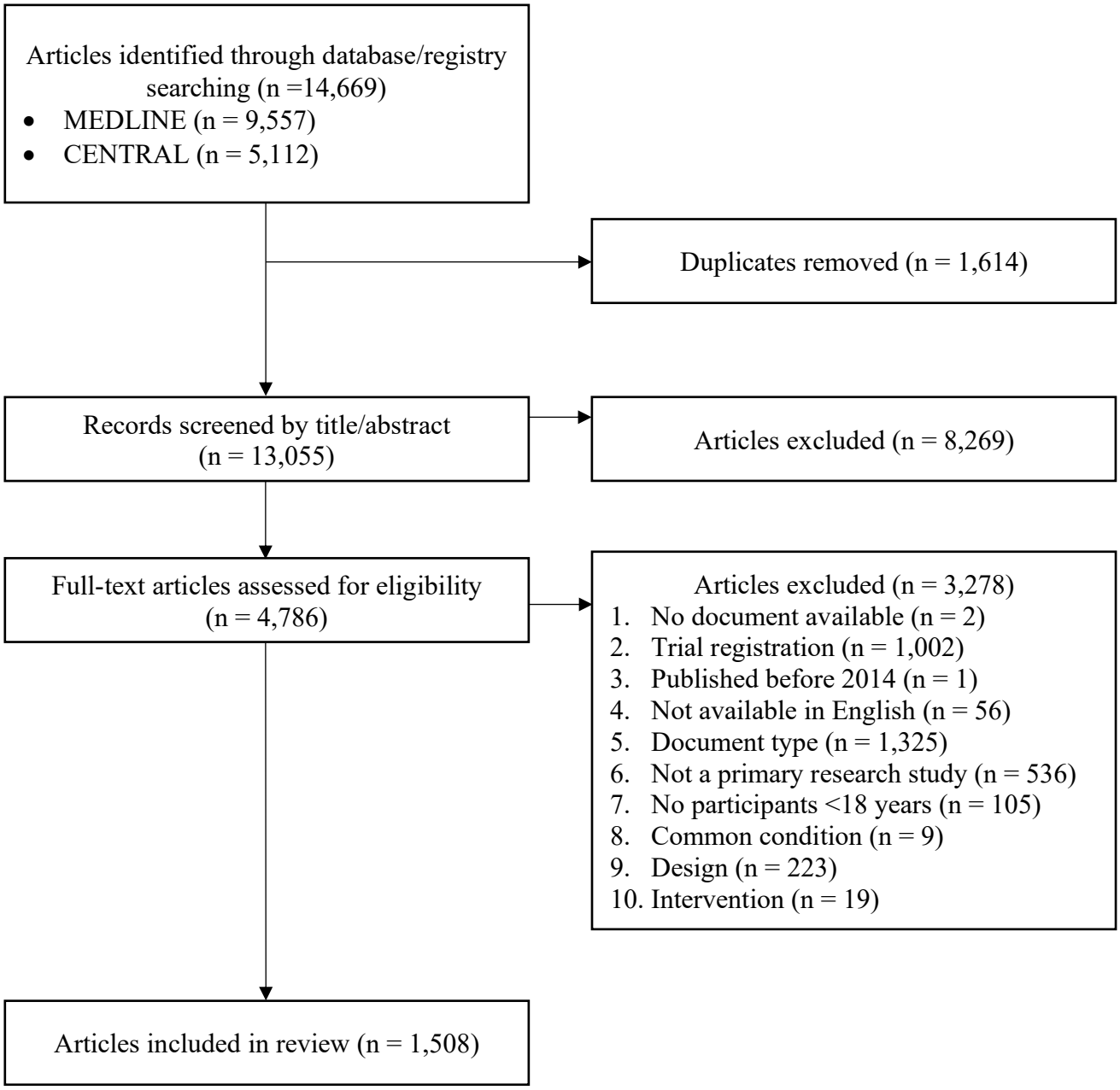
#15927 (Pignolo et al., 2023)[66]	Reduction of new heterotopic ossification (HO) in the open-label, phase 3 MOVE trial of palovarotene for fibrodysplasia ossificans progressiva (FOP)	None	Palovarotene	No other articles with the same source of external data.
Metachromatic leukodystrophy				
#14828 (Fumagalli et al., 2022)[62]	Lentiviral haematopoietic stem-cell gene therapy for early-onset metachromatic leukodystrophy: Long-term results from a non-randomised, open-label, phase 1/2 trial and expanded access	None	Atidarsagene autotemcel	No other articles with the same source of external data.
Mucopolysaccharidosis				
#7532 (Hendriksz et al., 2016)[52]	Long-term endurance and safety of elosulfase alfa enzyme replacement therapy in patients with morquio a syndrome	None	Elosulfase alfa	No other articles with the same source of external data.
Neuronal ceroid liposcinosis type 2				
#9629 (Schulz et al., 2018)[55]	Study of intraventricular cerliponase alfa for CLN2 disease	#9020 (final report of extension).	Cerliponase alfa	Both articles used data from the dementia in childhood neuronal ceroid lipofuscinosis patient database for their external control.
#9020 (Schulz et al., 2024)[69]	Safety and efficacy of cerliponase alfa in children with neuronal ceroid lipofuscinosis type 2 (CLN2 disease): An open-label extension study	#9629 (primary report)	Cerliponase alfa	
Spinal muscular atrophy				
#3403 (Krosschell et al., 2018)[54]	Clinical trial of l-carnitine and valproic acid in spinal muscular atrophy type 1	None	Levocarnitine and Valproic acid	No other articles with the same source of external data.
#83 (Day et al., 2021)[59]	Onasemnogene abeparvovec gene therapy for symptomatic infantile-onset spinal muscular atrophy in patients with two copies of SMN2 (STR1VE): An open-label, single-arm, multicentre, phase 3 trial	None	Onasemnogene abeparvovec	No other articles with the same source of external data.
#15926 (Mercuri et al., 2023)[65]	Risdiplam in types 2 and 3 spinal muscular atrophy: A randomised, placebo-controlled, dose--finding trial followed by 24 months of treatment	None	Risdiplam	No other articles with the same source of external data.
#8217 (Muntoni et al., 2020)[58]	Long-term follow-up of patients with type 2 and non-ambulant type 3 spinal muscular atrophy (SMA) treated with olesoxime in the OLEOS trial	None	Olesoxime	No other articles with the same source of external data.

eTable 4. Analytic methods used to address imbalances across the internal and external arms and the covariates considered

Article	Analytic method(s)	Baseline covariates
(Konstan et al., 2017)[53]	Propensity score matching	Age, sex, race, height-for-age, weight-for-age, body mass index, cystic fibrosis-related diabetes, percent predicted forced expiratory volume in one second (ppFEV1), percent predicted forced vital capacity (ppFVC), percent predicted FEV1/FVC ratio, percent predicted forced expiratory flow (midexpiratory range), and decile for baseline ppFEV1, baseline tests for Burkholderia species, aspergillys, alcaligenes, stenotrophomonas, pseudomonas aeruginosa cultures, methicillin-resistant staphylococcus aureus, and methicillin-sensitive staphylococcus aureus and use at baseline of acetylcysteine or mucomist, inhaled aztreonam, colistin, dornase alfa, leukotriene modifiers, and tobramycin.
(Clemens et al., 2020)[57]	Group matching	Age, functional status at baseline, geographic location, and glucocorticoid treatment status.
(Clemens et al., 2022)[61]	Group matching and Regression analysis conditioning on specific covariates	Group matching: Age, geographic location, ambulatory status and glucocorticoid treatment status. Regression analysis: Age and baseline measure of outcome.
(Clemens et al., 2023)[64]	Group matching	Age, functional status at baseline, geographic location, and glucocorticoid treatment status.
(Harper et al., 2024)[68]	Propensity score matching	Age, height, weight, ambulatory status, glucocorticoid treatment status and parameters of forced vital capacity.
(Khan et al., 2019)[56]	Regression analysis conditioning on specific covariates	Age and height.
(McDonald et al., 2021)[60]	Individual matching and Regression analysis conditioning on specific covariates	Matching: Genotype and "baseline characteristics" (did not specify further). Regression: Age.
(Mah et al., 2022)[63]	Individual matching	Age at baseline, baseline time to stand in seconds and baseline time to run/walk 10 meters in seconds.
(Zaidman et al., 2023)[67]	Inverse probability of treatment weighting using propensity scores	Age group, baseline North Star Ambulatory Assessment total score and key timed function tests (supine to stand and 10-meter walk/run).

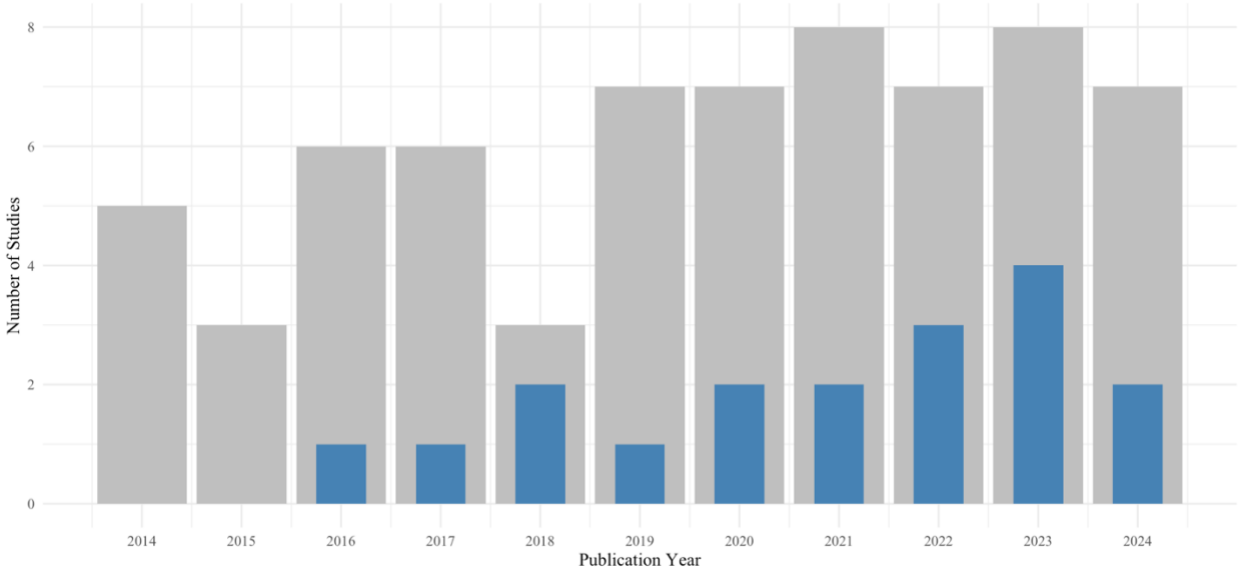
(Pignolo et al., 2023)[66]	Regression analysis conditioning on specific covariates	Age and sex.
(Fumagalli et al., 2022)[62]	Individual matching and Sibling matched	Age and disease subtype.
(Hendriksz et al., 2016)[52]	Regression analysis conditioning on specific covariates	Age, baseline 6-minute walk test score and baseline score of outcomes (3-minute stair climb test or normalized urine keratan sulfate).
(Schulz et al., 2018)[55]	Individual matching and Regression analysis conditioning on specific covariates	Matching: Age, genotype, sex and CLN2 clinical rating scale motor-language score. Regression for non-matched: Age, genotype, sex, baseline CLN2 clinical rating scale motor-language score, CLN2 clinical rating scale motor score and CLN2 clinical rating scale language score.
(Schulz et al., 2024)[69]	Individual matching and Regression analysis conditioning on specific covariates	Matching: Age, genotype and baseline CLN2 clinical rating scale motor-language score. Regression for non-matched: Age, genotype, sex and baseline CLN2 clinical rating scale motor-language score.
(Krosschell et al., 2018)[54]	Individual matching	Age and gender.
(Day et al., 2021)[59]	Individual matching	Age and gender.
(Mercuri et al., 2023)[65]	Inverse probability of treatment weighting using propensity scores	Age, sex, disease subtype, SMN2 copy number, scoliosis status and baseline total score for motor function measure.
(Muntoni et al., 2020)[58]	Individual matching and Regression analysis conditioning on specific covariates	Matching: Gender and disease subtype Regression: Age, gender, disease subtype and baseline motor function D1 + D2 scores.

Note. Covariates were extracted verbatim from study reports, protocols, statistical analysis plans or supplemental material. Variables such as "sex" and "gender" may have been used interchangeably across studies.



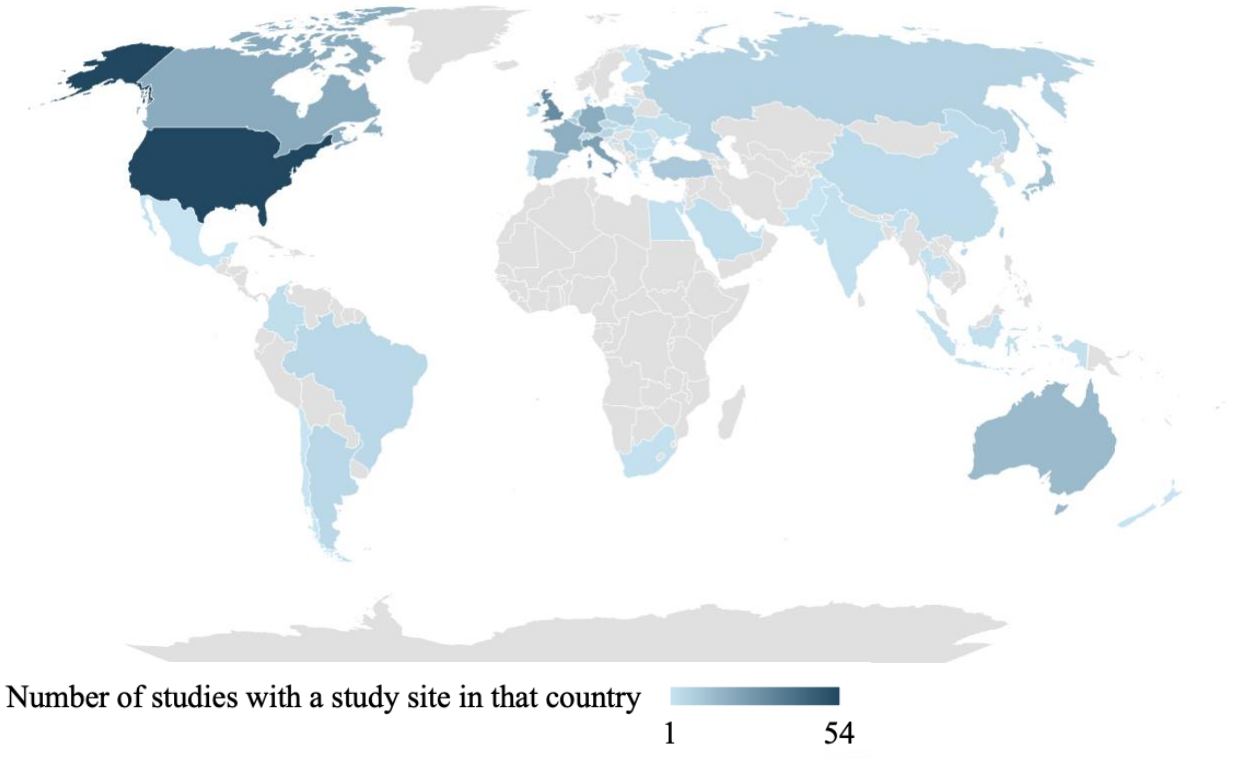
eFigure 1. PRISMA flow diagram depicting the identification, screening and selection of articles included in the RareKids-CAN living scoping review

Note. Reasons for exclusion are listed in order of exclusion. Document type: excluded studies included conference abstracts and proceedings, short summaries/abstracts that are not conference abstracts (lay summaries), commentaries, editorials, policy briefs, opinion pieces, corrections/errata, retractions, correspondence, letters to the editor or dissertations. Primary research study: involved the planned or completed collection of new data. Design: excluded studies included animal studies, in vitro studies, studies with no intervention, studies with retrospectively collected data, observational studies (case series, case studies, cohort studies, case-control studies and nested case-control studies) and post-market real-world data studies. Intervention: excluded studies whose main purpose was not to evaluate an intervention designed to improve health-related outcomes in the population. The number of studies included in the RareKids-CAN living inventory is up to date as of August 18th, 2026.



eFigure 2. Publication year of included articles

Note. The grey bands represent all included articles (n=67). The blue bands represent all articles that met all additional criteria and were prioritized for in-depth extraction (n=18).



eFigure 3. Global distribution of overall included studies' sites

Note. Study site location is missing for 1/67 articles.

Chapter 3: Preface

Objective: Summarize existing, publicly available guidance and criteria published by leading regulatory agencies and health technology assessment organizations, directed to sponsors/trialists or evaluators/reviewers of externally controlled pediatric rare disease trials and evaluate the alignment of published externally controlled pediatric genetic rare disease trials with this guidance.

Author contributions: Gabrielle Pratt Tremblay led the conception and design of the study, the screening, extraction, analysis for the literature review and interpretation of the data, and drafted the manuscript. Drs. Alex Bliu and Christopher Gravel contributed to the conception, design, and supervision of the study, interpretation of the data, and critically reviewed the manuscript. Dr. Beth Potter contributed to the conception, design, and supervision of the study, screening and verification of the extracted documents, analysis and interpretation of the data, and critically reviewed the manuscript.

Publication: This manuscript is prepared for submission to the Nature Communications Journal. Additional details in the methods and results have been added in this version to improve readability for the examiners of this thesis.

3. Manuscript 2 - *Guidance for externally controlled trials: A review and application to pediatric rare disease*

Gabrielle Pratt Tremblay[1], Alex Bliu[2], Christopher Gravel[1,3,4], Beth K. Potter[1]

Affiliations:

[1]School of Epidemiology and Public Health, University of Ottawa, ON, Canada,

[2]Office of Biostatistics, Biologic and Radiopharmaceutical Drugs Directorate, Health Canada, Ottawa, ON, Canada

[3]Department of Mathematics and Statistics, University of Ottawa, Ottawa, ON, Canada

[4]Data Literacy Research Institute, University of Ottawa, ON, Canada

Funding

This study is funded by the RareKids-CAN: Pediatric Rare Disease Clinical Trials and Treatment Network. RareKids-CAN is funded by the Canadian Institutes of Health Research as part of the Government of Canada's National Strategy of Drugs for Rare Diseases (Grant #192262).

3.1 Abstract

Externally controlled trials are frequently used in contexts where randomized controlled trials may be infeasible or unethical, such as in evaluations of new drugs for pediatric genetic rare diseases (RDs). Treatments evaluated using this design often face challenges in receiving regulatory approval and/or feasible reimbursement recommendations due to uncertainty about their evidence. We synthesized regulatory and health technology assessment guidance on this study design and assessed alignment of practices in externally controlled pediatric genetic RD trials against the guidance. We extracted and synthesized guidance from seven documents. Guidance was most often related to increasing comparability of participants across internal and external arms. Additional recommendations included considerations for outcome assessment, missing data and sensitivity analyses. Comparisons to externally controlled trials in pediatric genetic RD revealed an important gap in the rigor of current practices. More comprehensive guidance is needed to support trialists in the conduct of robust externally controlled trials.

3.2 Background

Rare diseases (RDs) are defined in the European Union and Canada as any disease that affects fewer than 1 in 2,000 people,[1,2] and in the United States as affecting fewer than 200,000 people.[3] RDs most often manifest during childhood[4] and often present significant barriers to daily living and reductions in life expectancy: an estimated 30% of children with RDs do not survive beyond 5 year of age.[4,5] It has been recently estimated that approximately 300 million people globally live with one or more RDs.[4,6] Despite their collective prevalence and significant impact, approximately 95% of RDs lack an approved treatment.[7]

The development of treatments for RDs is challenging due to the small population sizes, a lack of understanding of disease natural history, rapid progression for many RDs, and heterogeneity within diseases.[8,9] In response to the work of many RD patients, patient advocacy groups, and researchers,[10] several countries have recognized the high unmet need among this population and have launched legislative changes, special protocols, financial incentives and designated pathways to encourage research and development and promote accelerated access to disease-modifying therapies for RDs.[3,9,11-13] These opportunities in combination with recent advances in diagnostic technologies, biopharmaceuticals and genetic research, such as next-generation sequencing, enzyme replacement therapy, gene editing and gene therapies, have contributed to an increase in therapy development.[14,15] Despite these policies and technological advancements, important challenges remain. One such challenge relates to the feasibility of designing and conducting robust randomized controlled trials (RCTs) to establish the efficacy of new therapies. RCTs are considered the gold standard study design to establish treatment efficacy,[16] yet they are often infeasible in populations with RDs. Specifically, for RDs that do not have an existing efficacious treatment, establishing an internal control group, whereby a new and

potentially disease-modifying therapy is withheld from some participants, may be considered unethical, particularly for progressive and life-limiting conditions diagnosed in early childhood. In addition, RCTs can be impractical for RDs due to the small number of patients globally.[17-19]

In an attempt to address these feasibility and ethical concerns while providing robust evidence, an increasing number of trialists have begun conducting externally controlled trials, whereby outcomes from patients enrolled in a trial who receive a test treatment are compared to outcomes in individuals who are not enrolled in the trial and who did not receive the same treatment.[20] Many regulatory agencies and health technology assessment (HTA) bodies recognize the value of this study design in situations where an internal comparison group is not feasible or where unmet need is high, particularly in pediatric RDs.[20-24] However, they have also raised important concerns regarding the evidentiary robustness and interpretability of these trials. Treatments evaluated using externally controlled trials often face greater challenges in receiving regulatory approval and positive reimbursement recommendations, which stems from uncertainty surrounding the strength of the evidence generated by these designs, including concerns about data quality and potential biases, such as unmeasured confounding, selection bias, and information bias, particularly when external data sources may not be fully comparable to the associated single-arm clinical trial.[22,24,25]

Despite these persistent concerns, there remains limited guidance for clinical trialists on the appropriate design and conduct of externally controlled trials. It also remains unclear whether current practices in externally controlled trials for RDs align with existing guidance, or whether that guidance requires adaptation to address the specific challenges of RD research, particularly in pediatric settings where new therapies for genetic conditions are emerging rapidly. In an effort to support trialists, sponsors and real-world evidence producers in the design and conduct of

robust externally controlled trials, and to support evidence reviewers of such trials, we aimed to: 1) systematically summarize the existing regulatory and HTA guidance specific to externally controlled trials; 2) describe the alignment of this guidance with a risk-of-bias tool; and 3) explore the applicability of the guidance in the context of pediatric genetic RDs.

3.3 Methods

We conducted a focused review of regulatory and HTA guidance documents to identify guidance specific to externally controlled trials. In this review, an externally controlled trial was defined in accordance with the United States Food and Drug Administration's (FDA) definition which is a trial in which "outcomes in participants receiving the test treatment according to a protocol are compared to outcomes in a group of people external to the trial who had not received the same treatment".[20,p.1] Searches were conducted using the public website portals and guidance repositories of regulatory and HTA organizations.

3.3.1 *Eligibility criteria*

Documents were eligible for inclusion if they: 1) were written and published by a founding or standing regulatory member of the International Council for Harmonization of Technical Requirements for Pharmaceuticals for Human Use (ICH), including the ICH as an organization itself, or by an HTA organization that is a member of the International Network of Agencies for Health Technology Assessment (INAHTA) (see **Supplementary Methods 1** for a complete list of eligible organizations); 2) were a draft or final document directed to those who conduct trials and/or to those who review evidence; 3) contain recommendations or guidance on the design, analysis, reporting or examination of external controlled studies; and 4) were published in

English on or before June 16th, 2025. We considered documents that were specific to externally controlled trials and those that were not but that still contained one or more recommendations about these trials (e.g., guidance documents about real-world evidence that included sections about the use of such evidence in externally controlled trials).

If an eligible organization published multiple documents that contained recommendations or guidance about externally controlled trials, we first identified any documents that were solely and specifically focused on externally controlled trials from that organization. When such documents were available, they were selected for extraction and other documents from that organization were excluded on the assumption that the focused document(s) reflected the organization's primary advice about these trials. If there was no document specific to externally controlled trials, we included all documents from the organization that addressed the topic of externally controlled trials in part, as no single source could be considered the organization's authoritative recommendations.

3.3.2 *Search, screening and selection of sources of evidence*

We searched the website of each eligible regulatory and HTA organization for documents meeting all inclusion criteria on June 16th, 2025. See **Supplementary Methods 2** for the search terms used. One reviewer (lead author G.P.T) screened the title and executive summary as well as the full-text document against the predefined eligibility criteria. Full-text documents that met all inclusion criteria were verified by a second reviewer (B.K.P). We resolved any conflicts through discussion until a consensus was reached.

3.3.3 *Data extraction and coding*

Following full-text screening, one reviewer (G.P.T) independently extracted all verbatim guidance statements that pertained to the justification, design, conduct, analysis, reporting, and appraisal of externally controlled trials, from each document that met all inclusion criteria. Extraction was verified by a second reviewer (B.K.P), and we resolved any discrepancies through discussion until a consensus was reached.

We used qualitative content analysis to categorize and describe the topics covered by the verbatim guidance statements.[26] We first categorized each verbatim extract into one of three predefined categories: 1) criteria for the use of external controls; 2) design considerations; or 3) analysis considerations. Subsequently, we assigned one or more topics to each verbatim extract within each category. Topics were developed inductively to capture recurring, overlapping, or new concepts. We then summarized each verbatim extract into one or more key message to capture their core recommendations about the topics.

To synthesize the recommendations, we organized the verbatim extracts and associated key messages by category and topic. Extracts with overlapping topics and key messages were merged into a single synthesized guidance statement while preserving the original meaning as much as possible. Extracts representing unique messages or concepts were retained and rephrased to summarize information and increase consistency across all guidance statements. Coding and synthesis were conducted iteratively, returning to earlier steps as needed to develop a cohesive and consistent final list of guidance statements. All steps of content analysis were conducted by G.P.T and verified by B.K.P, with meetings throughout the process of iteration. Feedback from the full study team informed the final list of guidance statements.

3.3.4 *Comparison between the extracted guidance and a risk of bias tool*

To determine whether the available guidance from regulatory and HTA agencies aligns with existing risk-of-bias tools that may already be used by reviewers of externally controlled trials and by trialists or sponsors seeking to design such trials with mitigation of bias in mind, we mapped all revised guidance statements to the domains of the Risk Of Bias In Non-randomized Studies of Interventions version 2 (ROBINS-I V2) tool.[27] ROBINS-I V2 is a tool developed to assess the risk of bias from a non-randomized study that examined the effect of an intervention on an outcome and was thus deemed most suitable.

ROBINS-I V2 includes six domains, each of which aim to assess a specific risk of bias through unique signaling questions: risk of bias due to confounding; risk of bias in classification of interventions; risk of bias in the selection of participants into the study/analysis; risk of bias due to missing data; risk of bias arising from measurement of the outcome; and risk of bias in selection of the reported results. Following coding and synthesis, all final guidance statements were mapped to one or more domains of the ROBINS-I V2 tool by one reviewer (G.P.T) and verified by a second reviewer (B.K.P). Any discrepancies were resolved through discussion.

3.3.5 *Comparison to current practices*

We previously conducted a scoping review, in accordance with JBI methodology,[28] to identify and describe the design and conduct of externally controlled trials conducted in pediatric populations with genetic RDs, published between 2014 and 2025. More specifically, studies were included if they: 1) met the FDA definition of an externally controlled trial;[20] 2) ≥ 50% of study participants were diagnosed with a RD of genetic and non-oncologic origin, as indicated in the Orphanet Rare Disease Database and/or the National Institutes of Health Genetic and Rare

Diseases list;[29,30] and 3) the study reported comparisons of internal trial data to external data for one or more safety or efficacy outcome. The scoping review identified 67 articles that met all criteria for inclusion.[31-97] Of the included articles, 16 (24%) described trials that made a comparison to aggregate external data,[31-46] and 51 (76%) reported a comparison to individual-level external data.[47-97] Of these 51, a subset of 18 articles (27%) described trials that attempted to address differences in baseline characteristics between internal and external arms statistically, beyond adjustment for age and baseline outcome scores.[47-64] The scoping review focused specifically on the methodological and reporting characteristics of these 18 externally controlled trials (hereafter referred to as the "prioritized subset"). Findings from the scoping review were integrated into this study to identify areas of alignment and gaps between current practices specific to externally controlled pediatric genetic RD trials and existing regulatory and HTA guidance. Many of the trials from the scoping review were published prior to the publication of the guidance document. Therefore, our goal was not to determine whether trialists are adhering to existing guidance but rather to identify opportunities for improved practice in the pediatric genetic RD setting.

3.3.6 *Reporting and synthesis across study components*

A PRISMA flow diagram was used to summarize the document selection process, including the identification, screening and inclusion of documents with guidance statements pertaining to external controls. Summary tables were created to present the verbatim guidance extracts and their assigned categories, topics and key messages. We present tables of the final synthesized guidance statements and their mapping to ROBINS-I V2 domains, along with a narrative summary. For guidance statements related to the design and analysis considerations of externally

controlled trials, we also report the extent to which current practice in pediatric genetic RD trials aligned with guidance, using the information collected in our previously conducted scoping review.

3.4 Results

3.4.1 *Selection of documents containing guidance for the design, conduct and analysis of external controls*

Of the 60 eligible regulatory and HTA organizations, two did not have active websites and 16 did not have a website or documents available in English (**Figure 1**). Among the remaining 42 organizations, 37 did not have any available guidance documents that contained guidance specific to externally controlled trials or the use of external controls. Ultimately, two regulatory bodies (the FDA and ICH) and three HTA organizations (Canada's Drug Agency (CDA), the National Institute for Health and Care Excellence (NICE) and the French Authority for Health (HAS)), were found to have at least one document that met all inclusion criteria (**Table 1**). A total of seven documents were included in our extraction: one was a draft guidance document published by the FDA specifically focused on externally controlled trials,[20] and six were final guidance or methodological documents that included but were not specific to externally controlled trials, published by the ICH (n=2), CDA (n=1), HAS (n=2) and NICE (n=1).[98-103] Of the included documents, the earliest year of publication was 2000,[101] and the remaining documents were published between 2021 and 2025. No included documents were specific to RD or pediatric populations.

3.4.2 *Extraction and coding of guidance specific to external controls*

From the seven included documents, we extracted 70 verbatim statements (**Table 1 and SI 1**). The majority of extracts were taken from the FDA's *Considerations for the design and conduct of externally controlled trials for drug and biological products* (n=40),[20] followed by the ICH's *Guideline E10: Choice of control group and related issues in clinical trials* (n=11), the NICE's *Real-world evidence framework* (n=8),[103] and the HAS's *Methodology for the clinical development of medical devices* (n=7) .[98] The remaining documents included ≤2 extracted statements each.

Verbatim extracts were categorized as specific to criteria for using an external control (n=14), design considerations (n=43), and/or analysis considerations (n=18) for externally controlled trials (**Table SI 2**). The topics we developed iteratively to further classify extracts were the following: protocol/statistical analysis plan, disease, possibility of randomization, effect size, comparability across arms, outcomes, fit for purpose data, pooling data sources, analytic methods, missing data, and additional analyses. Eleven of the 70 verbatim extracts were mapped to two or more messages. Following the synthesis of similar messages, we created 33 final guidance statements, six of which were criteria for using an external control, 17 design considerations, and 10 analysis considerations (**Tables 2-6**).

Six of the 33 final guidance statements were categorized as criteria for using an externally controlled trial (**Table 2**). The remaining 27 guidance statements underwent mapping to ROBINS-I V2 domains: risk of bias due to confounding (n=7 guidance statements), outcome measurement (n=5 statements), selection of participants (n=1 statement), and missing data (n=3 statements) (**Tables 3-5**). The remaining guidance statements (n =11) were not specific to any one risk of bias but touched on several domains (**Table 6**).

3.4.3 *Criteria for the use of an external control*

Several reviewed documents included an overall recommendation that externally controlled trials are acceptable in cases where a randomized comparison is not feasible or ethical, or when direct comparative evidence is unavailable (**Table 2**).[98,100,101] External controls were thought to be most appropriate when disease progression is well understood and highly predictable,[20,101,102] a large treatment effect is anticipated,[20,101] external data sources include information on well-understood prognostic factors,[20,101] and the impact of baseline and treatment variables are well characterized and endpoints objectively measured.[101] In addition, treatment attributes (aside from the intervention evaluated in the trial) and related factors should be assessed a priori and deemed comparable across internal and external arms as a criterion for proceeding.[20]

3.4.4 *Externally controlled pediatric rare genetic disease trials*

From our scoping review of externally controlled pediatric genetic RD trials, our prioritized subset of articles included 18 articles reporting on trials that relied on individual-level external control data and applied analytic methods to address covariate imbalance beyond adjusting for age and baseline outcomes scores.[47-64] These 18 articles included participants across nine different RDs, including cystic fibrosis (n=1), Duchenne muscular dystrophy (n=8), fibrodysplasia ossificans progressiva (n=1), metachromatic leukodystrophy (n=1), mucopolysaccharidosis type 4A (n=1), neuronal ceroid lipofuscinosis type 2 (CLN2) (n=2), spinal muscular atrophy type 1 (n=3), spinal muscular atrophy type 2 (n=2) and spinal muscular atrophy type 3 (n=2). Sample sizes ranged from 10 to 516 in internal arms compared to the external control (median = 27) and from 4 to 1,588 participants in external arms (median = 65); aside from one larger study, all internal and external arms included fewer than 174 participants.

In the following paragraphs, we compare several characteristics of the studies described in these articles (**Table SI 3**) to the synthesized guidance (**Tables 3-6**).

3.4.5 *Guidance to minimize the risk of bias due to confounding*

Recommendations related to minimizing the risk of bias due to confounding in externally controlled trials were the subject of 7 guidance statements derived from 4 organizations. The FDA, ICH, HAS, NICE and CDA provided recommendations across four general topics: fit for purpose data, comparability across arms, analytic methods and additional analyses (**Table 3**).

First, where possible, organizations recommended that trialists select external control data with a source population similar to the internal arm, in an effort to minimize differences in key study variables such as patient characteristics, data collection, health care, geographic regions, and time periods.[20,100-103] In addition, trialists should rely on individual patient data from the external population,[101] including details on baseline characteristics and the study course. From the scoping review, all 18 articles reported on pediatric genetic RD trials that relied on individual-level data as this was a condition of inclusion in the prioritized subset (**Table SI 3**). Nonetheless, 16 of the 67 articles included in our overall scoping review relied on only aggregate external data.

To improve comparability across arms prior to analysis, regulators and HTA organizations recommended applying the same participant inclusion criteria across the internal and external arms.[20,103] All 18 articles included in our prioritized subset of trials in the scoping review were in alignment with this guidance (**Table 3; Table SI 3**). A more specific recommendation provided by the FDA is to compare prognostic factors across the internal and external arms, as they must be sufficiently similar to permit an unbiased assessment.[20] Fifteen

pediatric RD articles (83%) reported a comparison of baseline covariates between the internal and external arms. These comparisons were presented before applying a method to address imbalances (n=4; 27%), after applying such a method (n=9; 60%), or at both stages (n=2; 13%). Nonetheless, few of these articles discussed the comparability of the internal and external arms.

Additional analytic recommendations to improve the comparability across arms were: to specify an analytic method to identify and manage differences in baseline characteristics or other sources of confounding between the internal and external arms;[20,100] conduct additional analyses to evaluate and demonstrate the comparability of the internal and external arms on important factors, such as prognostic factors;[20] and conduct additional quantitative or qualitative bias analyses to assess the impact of confounding and other sources of bias.[20,103] As a condition of inclusion in the prioritized subset, all 18 articles in the scoping review applied a specific analytic method to address differences in baseline covariates, beyond adjusting for age or baseline outcome score. However, the majority of externally controlled trials for pediatric genetic RDs from our larger scoping review do not. Of the 18 articles, five discussed the success of their chosen analytic method to improve comparability, and none provided a bias analysis in their report.

3.4.6 *Guidance to minimize the risk of bias due to outcome measurement*

Guidance statements that were specific to minimizing the risk of bias due to outcome measurement were mainly drawn from the documents from the FDA and ICH. Specific recommendations related to the selection and measurement of outcomes were: to select objective and reliable outcomes;[20] to establish the timing and duration of outcome assessments in the internal trial arm prior to the selection of data for the external control;[20] and to ensure that the

timing and method of outcome assessment is consistent across the internal and external arms (**Table 4**).[20,101] In our prioritized subset of pediatric RD trials from the scoping review, all 18 articles reporting using the same outcome measurement instrument across internal and external arms but only 8/18 (44%) reported the same assessment time points across the internal and external arms for all outcomes that were compared (**Table SI 3**).

Other recommendations related to outcome measurement bias were specific to blinding. The FDA recommended that the study design and statistical analysis plan be blinded to results within external control data, with the exception of planned feasibility analyses,[20] and where possible, advised trialists/sponsors to consider blinding for outcome assessment in the internal and external arms.[20,101] Across the 18 articles included in the subset within our scoping review, only one reported blinding in the assessment or analysis of an outcome measured in the external control arm.

3.4.7 *Guidance to minimize the risk of bias due to the selection of participants*

Specific guidance regarding time-related biases was only found in the guidance document provided by the FDA.[20] This guidance emphasized the importance of considering immortal time bias when establishing an index date and follow-up period for outcome assessments in externally controlled trials and to ensure that these times are consistent and comparable across the internal and external arms (**Table 5**).[20] Among the prioritized subset of articles, 12 (67%) did not specify how the index date for the start of follow-up in their external arms was selected for at least one outcome (**Table SI 3**).

3.4.8 *Guidance to minimize the risk of bias due to missing data*

Guidance specific to the risk of bias due to missing data was only found in the FDA's document.[20] Prior to the conduct of the trial, it is advised to explore the extent of and reason for missing data in the external control data and consider the potential impact of missingness on the interpretation of study results (**Table 5**). Further to this, it is recommended to select an analytic method that includes a strategy to address missing data in external arm(s) and when analyzing the results of an externally controlled trial, the extent of missing data in both the internal and external arms should be assessed to examine their impact on study results. Of the prioritized subset of articles from our scoping review of pediatric genetic RD trials, four (22%) reported on missingness in baseline covariates in the internal arm(s), three (16%) of which also reported on missingness in baseline covariates in the external control(s), 11 (61%) reported on missingness in the outcome measurement of the internal arm(s) and nine (50%) reported on missingness in outcomes for the external arm(s) (**Table SI 3**). However, only one article clearly pre-specified how missingness in the external arm would be addressed.

3.4.9 *Guidance to minimize the risk of bias across several domains*

Several guidance statements were related to the risk of bias more broadly, mapping to more than one ROBINS-I V2 domain. The topics of these guidance statements included considerations for the protocol and statistical analysis plan, fit for purpose data, pooling of data sources, comparability across arms, the analytic method and additional analyses (**Table 6**).

Guidance statements that were general recommendations for externally controlled trial protocols included planning and specifying the selection of the external control data source, design elements and analyses prior to the conduct of the externally controlled trial, [20,99-102] and

avoiding revisions to the statistical analysis plan where possible.[20] More specific recommendations for the protocol and statistical analysis plan were to consider using target trial emulation and the estimand framework.[20,104,105] Of the prioritized subset of articles from our scoping review, 14 (78%) pre-specified the inclusion of an external control and 12 of the 13 (92%) with an accessible protocol pre-specified the external controlled design and analysis prior to the conduct of the trial (**Table SI 3**). No article reported using target trial emulation or the estimand framework in their available protocol or study report.

In the selection of data sources for the external control arm, some regulatory and HTA organizations recommended that, where possible, trialists should consider using data from a previous clinical trial.[20,103] However, only 2 articles (11%) from our prioritized subset of pediatric RD trials reported using data from a previously conducted clinical trial. In cases where no single optimal source of external control data exists, the ICH noted that it is acceptable to study multiple controls, if pre-planned and analyzed/interpreted conservatively.[101] If multiple potential sources of external data are suitable, pooling should only be done in cases where there is limited heterogeneity in terms of coverage and data quality.[103] Of the prioritized subset of pediatric genetic RD trials from the scoping review, 5/18 articles (28%) included two or more external control arms in their analysis. Of the 13 articles with a single external control arm, three pooled more than one data source to create the external arm: one study combined retrospective routinely collected data and prospective natural history data for the same participants,[53] a second pooled data from three unique sources, including a natural history study and two clinical trials,[62] and a third pooled data from two unique sources, including a natural history study and the placebo arm of a clinical trial.[60] Regardless of pooling, NICE recommended that any differences

in variable definitions and data collection across the internal and external arms should be reported.[103]

Finally, guidance statements about analysis that were relevant to multiple domains of ROBINS-I V2 included recommendations to: consider a conservative analytic approach, including the potential for a more stringent statistical significance level;[101] justify the analytic method selected for comparison of the internal trial arm(s) with the external control and examine its assumptions;[20] and use sensitivity analyses to evaluate the potential impact of assumptions, potential confounders, missing data and other sources of biases.[20,100] Within the prioritized subset of articles from our scoping review of pediatric RD trials, 5/18 articles (28%) included a sensitivity analysis related to the external control, each of which was used to evaluate the potential impact of confounders, missing data, or other potential sources of bias.

3.5 Discussion

3.5.1 *Summary of guidance*

To our knowledge, this literature review provides the first summary of publicly available regulatory and HTA guidance for the design and conduct of externally controlled trials. From our synthesis across organizations, key recommendations for trialists and sponsors to consider when planning and conducting an externally controlled trial were to: 1) plan and report the characteristics of the external controls, design elements and analyses prior to the conduct of an externally controlled trial; 2) increase comparability across the internal and external arms at baseline by selecting an external control that is similar to the internal arm(s) in patient characteristics, geographic distribution, data collection and time periods, and applying analytic methods to address any confounding variables; 3) increase the comparability of outcome

assessments by selecting objective outcomes, considering blinding in the outcome assessment or analysis of both the internal and external data; 4) select an index date to start follow up and a follow up time period that is comparable across arms; and 5) conduct sensitivity analyses to evaluate the potential impact of confounders, missing data or other potential sources of bias and to further demonstrate the rigor of the study findings.

Several of the guidance statements presented in this review may also be beneficial for researchers and registry holders designing data collection for registries, natural history studies, and cohort studies, that could serve as a source of external control data for clinical trials. For example, where possible, data producers should collect a comprehensive set of prognostic factors and baseline variables, select and measure objective outcomes at protocolized time points, and prevent and explore reasons for missing data. Taking these synthesized guidance statements into consideration may increase the reliability of these data sources and improve the robustness of externally controlled trials that incorporate them.

3.5.2 *Alignment of guidance across organizations*

Many guidance statements extracted in this review were found to be concordant across organizations, with no conflicting advice. However, the quantity and comprehensiveness of the guidance varied widely across documents and organizations. Except for the FDA,[20] the guidance available for external controls was often a small section in a larger document focused on real-world data or clinical trial design. This is expected, given that externally controlled trials both implement a clinical intervention study and incorporate observational data. Nonetheless, future guidance on externally controlled trials should aim for comprehensiveness and include more specific recommendations, as this may increase their value to trialists and evidence reviewers.

In the evaluation of externally controlled trials, regulatory agencies and HTA organizations often raise similar concerns.[22] However, given the differences in aims between regulatory and HTA bodies, flaws in the conduct of externally controlled trials may have different impacts. For example, treatments submitted with weak comparative evidence and high uncertainty may be given a conditional approval by a regulatory agency based on a high unmet need but given a negative reimbursement recommendation by an HTA organization. In order to support decision making for both regulatory agencies and HTA organizations, as well as increase accessibility of potentially disease-modifying therapies for patients, harmonized guidance on the conduct of robust externally controlled trials is needed from these organizations.

3.5.3 *Alignment of guidance with ROBINS-I V2*

We demonstrated that certain domains of ROBINS-I V2 may align well with the guidance provided by the regulatory and HTA organizations. Approximately 40% of the synthesized guidance statements that were specific to design or analytic considerations were not specific to one domain but instead simultaneously addressed several domains of bias from ROBINS-I V2. Where guidance statements were domain-specific, they most commonly related to limiting the risk of bias due to confounding, highlighting the particular importance of confounding to externally controlled trials. Other domains that were represented in the guidance included risk of bias due to the measurement of the outcome, selection of participants, or missing data; guidance statements specific to these three sources of bias were more limited and primarily taken from the FDA. Despite their low representation in the current guidance, these sources of bias are important concerns when incorporating external sources of data that are not specifically collected for that clinical intervention study. No guidance was specific to the other two domains of ROBINS-I V2 (bias in the classification of interventions and bias in the selection of reported

results). In addition to more detailed guidance, we recommend the development of a risk-of-bias tool that is specific to externally controlled trials. Given the unique challenges and sources of bias associated with this design, such a tool may be beneficial for both trialists and reviewers of externally controlled trials (e.g., for regulatory or HTA purposes). Pending the publication of more detailed guidance and a specific risk-of-bias tool, trialists may consider the guidance we synthesized alongside ROBINS-I V2 in the design and reporting of their externally controlled trials in an effort to reduce the risk of bias in their results.

3.5.4 *Current practices in pediatric genetic RD*

In this review, we demonstrated that there are many opportunities for better practice among externally controlled trials in pediatric genetic RDs. Notable gaps between the current guidance and reported practices in pediatric genetic RD studies include the application of a specific analytic method to address balance in baseline characteristics and confounders across arms, the clear establishment of a comparable index date to initiate follow-up and alignment of follow-up periods across arms, blinding in the analysis of outcome measures and the reporting of additional analyses to address concerns of missingness and other sources of bias. These gaps in alignment with guidance must however be considered within their historical context: at the time of publication, for approximately half (55%) of the included articles, the only available guidance was the ICH's E10 guidelines.[101] While these articles are not necessarily failing to adhere to existing guidance (since much of the guidance was published recently), the gaps in the design and conduct that we identified nevertheless indicate a need for better practice. These concerns are not unique to externally controlled trials conducted in pediatric populations with RDs, as they have been raised in literature reviews conducted in other disease areas.[106,107]

Thoughtful consideration and application of the guidance presented in this review, together with transparent reporting, will be an important step to support the conduct of robust externally controlled trials in RD, as well as other populations that meet the above-mentioned criteria for the use of an external control.

3.5.5 *Estimand framework and target trial emulation*

Target trial emulation and the estimand framework[105,108] are important and complementary causal inference frameworks to consider in the design of externally controlled trials.[20,103] Well-designed RCTs are used to answer specific causal questions because, in addition to randomization, they explicitly specify key elements of the trial such as the eligibility criteria, treatment strategies, assignments, outcomes of interest, follow-up, and statistical analysis.[104,108,109] Together, these elements help inform and define the causal estimand.

Emulating a target RCT could be a beneficial strategy for designing an externally controlled trial, as it may lead to the incorporation of approaches that mitigate design biases common in observational studies, such as selection bias and immortal time bias.[104] It may also inform whether or not an externally controlled trial will be able to estimate the desired treatment effect. Ensuring the pre-specification of key design elements and subsequently, a well-defined estimand, will allow for increased interpretability and reliability of the study estimates. The importance of the implementation of these frameworks in externally controlled trials has been highlighted by Hogervorst et al.,[110] in their recent systematic review on analytic methods for comparing uncontrolled trials with external controls using real-world data.

3.5.6 *Limitations*

This review has some limitations that should be acknowledged. First, our review did not include guidance documents published by academic research groups or pharmaceutical companies. Although such groups may have produced guidance relevant to external controls, we limited our scope to regulatory and HTA organizations, as they are responsible for decisions regarding treatment approvals and reimbursements and thus have the ability to produce directive guidance that will be used to review the evidence generated by externally controlled trials. Second, this review was limited to publicly available documents published in English. Some eligible organizations may have internal documents relevant to external controls or provide confidential scientific advice, however, we considered publicly accessible guidance to be the most relevant source for trialists and sponsors, as they are most readily accessible to them. Furthermore, relevant documents may exist in other languages, however, translation of documents was not feasible in this study. Third, to identify eligible documents, we performed website searches, and it is possible that some relevant documents were missed due to website indexing, terminology differences across jurisdictions, or document accessibility issues. Fourth, when making our comparison to previously published externally controlled trials in pediatric genetic RD, the information collected was subject to author reporting. Consequently, some studies may have aligned more with guidance than what was observed but failed to report this in their publication. This review did not compare the views of different regulatory and HTA agencies regarding externally controlled trials. This would not have been possible given the varying quantity and comprehensiveness of guidance across documents. For example, the FDA provided an entire guidance document specific to externally controlled trials, whereas all other organizations published documents that ranged from one to eleven guidance statements specific to externally

controlled trials. Finally, qualitative content analysis involves subjective judgements, which introduces the risk of error. In an effort to limit this risk, all extraction, coding and synthesis was verified by a second reviewer.

3.5.7 *Conclusions*

In this review, we identified and synthesized publicly available guidance on the design, conduct and analysis of externally controlled trials from major regulatory bodies and HTA organizations. This synthesis offers a comprehensive resource for the planning and design of externally controlled trials that aligns with current regulatory and HTA expectations in jurisdictions that follow or adopt the guidance and recommendations of the CDA, FDA, HAS, ICH and NICE. Thoughtfully incorporating this guidance into the design, conduct and reporting of future externally controlled trials is an essential step to increase the rigor and acceptability of the evidence they generate. More guidance with specific recommendations aimed at addressing the risk of bias due to the selection of outcomes, the prevention of time-related biases and missing data is needed. Furthermore, we recommend that a risk-of-bias tool specific to externally controlled trials be developed as a complement to this guidance to support the robust conduct and evaluation of these trials.

3.6 References for Chapter 3

1. Canadian Organization for Rare Diseases. About CORD: Key facts. Accessed February 20, 2025. https://www.raredisorders.ca/about/about-cord

2. European Commission. Rare diseases: EU research on rare diseases. Accessed April 1, 2025. https://research-and-innovation.ec.europa.eu/research-area/health/rare-diseases_en

3. Orphan Drug Act, Pub. L. No. 97-414, 96 2049-2066 (1983). Accessed April 18, 2025. https://www.govinfo.gov/content/pkg/STATUTE-96/pdf/STATUTE-96-Pg2049.pdf

4. The landscape for rare diseases in 2024. *Lancet Glob Health*. 2024;12(3):e341. doi:10.1016/S2214-109X(24)00056-1

5. Dumbuya JS, Zeng C, Deng L, et al. The impact of rare diseases on the quality of life in paediatric patients: Current status. *Front Public Health*. 2025;13:1531583. doi:10.3389/fpubh.2025.1531583

6. Nguengang Wakap S, Lambert DM, Olry A, et al. Estimating cumulative point prevalence of rare diseases: Analysis of the Orphanet database. *Eur J Hum Genet*. 2020;28(2):165-173. doi:10.1038/s41431-019-0508-0

7. Fermaglich LJ, Miller KL. A comprehensive study of the rare diseases and conditions targeted by orphan drug designations and approvals over the forty years of the Orphan Drug Act. *Orphanet J Rare Dis*. 2023;18(1):163. doi:10.1186/s13023-023-02790-7

8. Bouwman ML, Sousa JJS, Pina MET. Regulatory issues for orphan medicines: A review. *Health Policy Technol*. 2020;9(1):115-121. doi:10.1016/j.hlpt.2019.11.008

9. Chen J, Nie L, Lee S, et al. Challenges and possible strategies to address them in rare disease drug development: A statistical perspective. *Clin Pharmacol Ther*. 2025;118(1):62-73. doi:10.1002/cpt.3631

10. Saltonstall P, Ross H, Kim PT. The Orphan Drug Act at 40: Legislative Triumph and the Challenges of Success. *Milbank Q*. 2024;102(1):83-96. doi:10.1111/1468-0009.12680

11. European Parliament, Council of the European Union. *Regulation (EC) No 141/2000 of 16 December 1999 on orphan medicinal products*. 2000:1-15. *OJL18*. Accessed April 18, 2026. http://data.europa.eu/eli/reg/2000/141/oj

12. Health Canada. Canada's regulatory approach to drugs for rare diseases: Orphan drugs. Government of Canada. Updated April 1, 2022. Accessed April 18, 2026. https://www.canada.ca/en/health-canada/services/licences-authorizations-registrations-drug-health-products/regulatory-approach-drugs-rare-diseases.html

13. Gammie T, Lu CY, Babar ZU. Access to orphan drugs: A comprehensive review of legislations, regulations and policies in 35 countries. *PLoS One*. 2015;10(10):e0140002. doi:10.1371/journal.pone.0140002

14. Ezgu FS. New therapeutic modalities for rare diseases. *Glob Pediatr*. 2024;7:100134. doi:10.1016/j.gpeds.2024.100134

15. Haffner ME. History of orphan drug regulation-United States and beyond. *Clin Pharmacol Ther*. 2016;100(4):342-343. doi:10.1002/cpt.426

16. Schulz KF, Grimes DA. Blinding in randomised trials: Hiding who got what. *Lancet*. 2002;359(9307):696-700. doi:10.1016/S0140-6736(02)07816-9

17. Khachatryan A, Read SH, Madison T. External control arms for rare diseases: Building a body of supporting evidence. *J Pharm Dyn*. 2023;50(6):501-506. doi:10.1007/s10928-023-09858-8

18. Augustine EF, Adams HR, Mink JW. Clinical trials in rare disease: Challenges and opportunities. *J Child Neurol*. 2013;28(9):1142-50. doi:10.1177/0883073813495959

19. Jahanshahi M, Gregg K, Davis G, et al. The use of external controls in FDA regulatory decision making. *Ther Innov Regul Sci*. 2021;55(5):1019-1035. doi:10.1007/s43441-021-00302-y

20. United States Food and Drug Administration. *Considerations for the design and conduct of externally controlled trials for drug and biological products guidance for industry: Draft guidance*. 2023. Accessed October 24, 2024. https://www.fda.gov/vaccines-blood-biologics/guidance-compliance-regulatory-information-biologics/biologics-guidances

21. Weberpals J, Wang SV. External controls to study treatment effects in rare diseases: Challenges and future directions. *Clin Pharmacol Ther*. 2024;116(6):1521-1524. doi:10.1002/cpt.3443

22. Sola-Morales O, Curtis LH, Heidt J, et al. Effectively leveraging RWD for external controls: A systematic literature review of regulatory and HTA decisions. *Clin Pharmacol Ther*. 2023;114:325-355. doi:10.1002/cpt.2914

23. Vaghela S, Tanni KA, Banerjee G, Sikirica V. A systematic review of real-world evidence (RWE) supportive of new drug and biologic license application approvals in rare diseases. *Orphanet J Rare Dis*. 2024;19(1):117. doi:10.1186/s13023-024-03111-2

24. Velummailum RR, McKibbon C, Brenner DR, Stringer EA, Ekstrom L, Dron L. Data challenges for externally controlled trials: Viewpoint. *J Med Internet Res*. 2023;25:e43484. doi:10.2196/43484

25. Rath A, Salamon V, Peixoto S, et al. A systematic literature review of evidence-based clinical practice for rare diseases: What are the perceived and real barriers for improving the

evidence and how can they be overcome? *Trials*. 2017;18(1):556. doi:10.1186/s13063-017-2287-7

26. Hsieh HF, Shannon SE. Three approaches to qualitative content analysis. *Qual Health Res*. 2005;15(9):1277-88. doi:10.1177/1049732305276687

27. Sterne J, Brandt Mathur M, Elbers R, et al. The Risk Of Bias In Non-randomized Studies – of Interventions, Version 2 (ROBINS-I V2) - Draft. Accessed April 1, 2026. https://drive.google.com/file/d/1tXCLVa7aFVaiV3OnitZDgAnX3kXWXor0/view

28. Peters MDJ, Marnie C, Tricco AC, et al. Updated methodological guidance for the conduct of scoping reviews. *JBI Evid Implement*. 2021;19(1):3-10. doi:10.1097/xeb.0000000000000277

29. Orphanet. Diseases. Accessed January 20, 2025, https://www.orpha.net/en/disease

30. National Institutes of Health. Genetic and Rare Diseases (GARD) Information Center: Diseases. Accessed January 20, 2025. https://rarediseases.info.nih.gov

31. Auerswald G, Kurnik K, Aledort LM, et al. The EPIC study: A lesson to learn. *Haemophilia*. 2015;21(5):622-628. doi:10.1111/hae.12666

32. Chien Y-H, Lee N-C, Tseng S-H, et al. Efficacy and safety of AAV2 gene therapy in children with aromatic L-amino acid decarboxylase deficiency: An open-label, phase 1/2 trial. *Lancet Child Adolesc Health*. 2017;1(4):265-273. doi:10.1016/S2352-4642(17)30125-6

33. Dvorak CC, Horn BN, Puck JM, et al. A trial of alemtuzumab adjunctive therapy in allogeneic hematopoietic cell transplantation with minimal conditioning for severe combined immunodeficiency. *Pediatr Transplant*. 2014;18(6):609–616. doi:10.1111/petr.12310

34. Gilchrist FJ, Bui S, Gartner S, et al. ALPINE2: Efficacy and safety of 14-day vs 28-day inhaled aztreonam for Pa eradication in children with cystic fibrosis. *J Cyst Fibros*. 2024;23(1):80-86. doi:10.1016/j.jcf.2023.06.008

35. Hankins JS, Aygun B, Nottage K, et al. From infancy to adolescence: Fifteen years of continuous treatment with hydroxyurea in sickle cell anemia. *Medicine*. 2014;93(28):e215. doi:10.1097/MD.0000000000000215

36. Hoppe JE, Kasi AS, Pittman JE, et al. Vanzacaftor-tezacaftor-deutivacaftor for children aged 6-11 years with cystic fibrosis (RIDGELINE Trial VX21-121-105): An analysis from a single-arm, phase 3 trial. *Lancet Respir Med*. 2024;13(3):244-255. doi:10.1016/S2213-2600(24)00407-7

37. Kavakli K, Yang R, Rusen L, Beckmann H, Tseneklidou-Stoeter D, Maas Enriquez M. Prophylaxis vs. on-demand treatment with BAY 81-8973, a full-length plasma protein-free recombinant factor VIII product: Results from a randomized trial (LEOPOLD II). *J Thromb Haemost*. 2015;13(3):360-369. doi:10.1111/jth.12828

38. Lynch DR, Farmer J, Hauser L, et al. Safety, pharmacodynamics, and potential benefit of omaveloxolone in Friedreich ataxia. *Ann Clin Transl Neurol*. 2019;6(1):15-26. doi:10.1002/acn3.660

39. Lynch DR, Hauser L, McCormick A, et al. Randomized, double-blind, placebo-controlled study of interferon-γ 1b in Friedreich Ataxia. *Ann Clin Transl Neurol*. 2019;6(3):546-553. doi:10.1002/acn3.731

40. Manco-Johnson MJ, Bomgaars L, Palascak J, et al. Efficacy and safety of protein C concentrate to treat purpura fulminans and thromboembolic events in severe congenital protein C deficiency. *Thromb Haemost*. 2016;116(1):58-68. doi:10.1160/th15-10-0786

41. Mendell JR, Al-Zaidy S, Shell R, et al. Single-dose gene-replacement therapy for spinal muscular atrophy. *N Engl J Med*. 2017;377(18):1713-1722. doi:10.1056/NEJMoa1706198

42. Mendell JR, Sahenk Z, Lehman K, et al. Assessment of systemic delivery of rAAVrh74.MHCK7.micro-dystrophin in children with Duchenne muscular dystrophy: A nonrandomized controlled trial. *JAMA Neurol*. 2020;77(9):1122-1131. doi:10.1001/jamaneurol.2020.1484

43. Patterson M, Ramaswami U, Donald A, et al. Disease-modifying, neuroprotective effect of N-acetyl-L-leucine in adult and pediatric patients with Niemann-Pick disease type C. *Neurology*. 2024;105(1):e213589. doi:10.1101/2024.10.11.24315318

44. Saxena K, Lalezari S, Oldenburg J, et al. Efficacy and safety of BAY 81-8973, a full-length recombinant factor VIII: Results from the LEOPOLD I trial. *Haemophilia*. 2016;22(5):706-712. doi:10.1111/hae.12952

45. Taylor M, Jefferies J, Byrne B, et al. Cardiac and skeletal muscle effects in the randomized HOPE-Duchenne trial. *Neurology*. 2019;92(8):e866-e878. doi:10.1212/WNL.0000000000006950

46. Windyga J, Lissitchkov T, Stasyshyn O, et al. Pharmacokinetics, efficacy and safety of BAX326, a novel recombinant factor IX: A prospective, controlled, multicentre phase I/III study in previously treated patients with severe (FIX level <1%) or moderately severe (FIX level ≤2%) haemophilia B. *Haemophilia*. 2014;20(1):15-24. doi:10.1111/hae.12228

47. Hendriksz CJ, Parini R, AlSayed MD, et al. Long-term endurance and safety of elosulfase alfa enzyme replacement therapy in patients with Morquio A syndrome. *Mol Genet Metab*. 2016;119(1-2):131-143. doi:10.1016/j.ymgme.2016.05.018

48. Konstan MW, McKone EF, Moss RB, et al. Assessment of safety and efficacy of long-term treatment with combination lumacaftor and ivacaftor therapy in patients with cystic fibrosis homozygous for the F508del-CFTR mutation (PROGRESS): A phase 3, extension study. *Lancet Respir Med*. 2017;5(2):107-118. doi:10.1016/S2213-2600(16)30427-1

49. Krosschell KJ, Kissel JT, Townsend EL, et al. Clinical trial of L-carnitine and valproic acid in spinal muscular atrophy type I. *Muscle Nerve*. 2018;57(2):193-199. doi:10.1002/mus.25776

50. Schulz A, Ajayi T, Specchio N, et al. Study of intraventricular cerliponase alfa for CLN2 disease. *N Engl J Med*. 2018;378(20):1898-1907. doi:10.1056/NEJMoa1712649

51. Khan N, Eliopoulos H, Han L, et al. Eteplirsen treatment attenuates respiratory decline in ambulatory and non-ambulatory patients with Duchenne muscular dystrophy. *J Neuromuscul Dis*. 2019;6(2):213-225. doi:10.3233/JND-180351

52. Clemens PR, Rao VK, Connolly AM, et al. Safety, tolerability, and efficacy of viltolarsen in boys with Duchenne muscular dystrophy amenable to exon 53 skipping: A phase 2 randomized clinical trial. *JAMA Neurol*. 2020;77(8):982-991. doi:10.1001/jamaneurol.2020.1264

53. Muntoni F, Bertini E, Comi G, et al. Long-term follow-up of patients with type 2 and non-ambulant type 3 spinal muscular atrophy (SMA) treated with olesoxime in the OLEOS trial. *Neuromuscul Disord*. 2020;30(12):959-969. doi:10.1016/j.nmd.2020.10.008

54. Day JW, Finkel RS, Chiriboga CA, et al. Onasemnogene abeparvovec gene therapy for symptomatic infantile-onset spinal muscular atrophy in patients with two copies of SMN2 (STR1VE): An open-label, single-arm, multicentre, phase 3 trial. *Lancet Neurol*. 2021;20(4):4284-4293. doi:10.1016/S1474-4422(21)00001-6

55. McDonald CM, Shieh PB, Abdel-Hamid HZ, et al. Open-label evaluation of eteplirsen in patients with Duchenne muscular dystrophy amenable to exon 51 skipping: PROMOVI trial. *J Neuromuscul Dis*. 2021;8(6):989-1001. doi:10.3233/JND-210643

56. Clemens PR, Rao VK, Connolly AM, et al. Long-term functional efficacy and safety of viltolarsen in patients with Duchenne muscular dystrophy. *J Neuromuscul Dis*. 2022;9(4):493-501. doi:10.3233/JND-220811

57. Fumagalli F, Calbi V, Natali Sora MG, et al. Lentiviral haematopoietic stem-cell gene therapy for early-onset metachromatic leukodystrophy: Long-term results from a non-randomised, open-label, phase 1/2 trial and expanded access. *Lancet*. 2022;399(10322):372-383. doi:10.1016/S0140-6736(21)02017-1

58. Mah JK, Clemens PR, Guglieri M, et al. Efficacy and safety of vamorolone in Duchenne muscular dystrophy: A 30-month nonrandomized controlled open-label extension trial. *JAMA Netw Open*. 2022;5(5):e2144178. doi:10.1001/jamanetworkopen.2021.44178

59. Clemens PR, Rao VK, Connolly AM, et al. Efficacy and safety of viltolarsen in boys with Duchenne muscular dystrophy: Results from the phase 2, open-label, 4-year extension study. *J Neuromuscul Dis*. 2023;10(3):439-447. doi:10.3233/JND-221656

60. Mercuri E, Baranello G, Boespflug-Tanguy O, et al. Risdiplam in types 2 and 3 spinal muscular atrophy: A randomised, placebo-controlled, dose-finding trial followed by 24 months of treatment. *Eur J Neurol*. 2023;30(7):1945-1956. doi:10.1111/ene.15499

61. Pignolo RJ, Hsiao EC, Al Mukaddam M, et al. Reduction of new heterotopic ossification (HO) in the open-label, phase 3 MOVE trial of palovarotene for fibrodysplasia ossificans progressiva (FOP). *J Bone Miner Res*. 2023;38(3):381-394. doi:10.1002/jbmr.4762

62. Zaidman CM, Proud CM, McDonald CM, et al. Delandistrogene moxeparvovec gene therapy in ambulatory patients (aged ≥4 to <8 Years) with Duchenne muscular dystrophy: 1-year interim results from study SRP-9001-103 (ENDEAVOR). *Ann Neurol*. 2023;94(5):955-968. doi:10.1002/ana.26755

63. Harper AD, Topaloglu H, Mercuri E, et al. Safety and efficacy of viltolarsen in ambulatory and nonambulatory males with Duchenne muscular dystrophy. *Sci Rep*. 2024;14(1):23488. doi:10.1038/s41598-024-70783-y

64. Schulz A, Specchio N, de Los Reyes E, et al. Safety and efficacy of cerliponase alfa in children with neuronal ceroid lipofuscinosis type 2 (CLN2 disease): An open-label extension study. *Lancet Neurol*. 2024;23(1):60-70. doi:10.1016/S1474-4422(23)00384-8

65. Hacein-Bey-Abina S, Pai S-Y, Gaspar HB, et al. A modified gamma-retrovirus vector for X-linked severe combined immunodeficiency. *N Engl J Med*. 2014;371(15):1407-1417. doi:10.1056/NEJMoa1404588

66. Levin SW, Baker EH, Zein WM, et al. Oral cysteamine bitartrate and N-acetylcysteine for patients with infantile neuronal ceroid lipofuscinosis: A pilot study. *Lancet Neurol*. 2014;13(8):777-787. doi:10.1016/S1474-4422(14)70142-5

67. Jones SA, Bialer M, Parini R, et al. Safety and clinical activity of elosulfase alfa in pediatric patients with Morquio A syndrome (mucopolysaccharidosis IVA) less than 5 y. *Pediatr Res*. 2015;78(6):717-722. doi:10.1038/pr.2015.169

68. Finkel RS, Chiriboga CA, Vajsar J, et al. Treatment of infantile-onset spinal muscular atrophy with nusinersen: A phase 2, open-label, dose-escalation study. *Lancet*. 2016;388(10063):3017-3026. doi:10.1016/S0140-6736(16)31408-8

69. Whyte MP, Madson KL, Phillips D, et al. Asfotase alfa therapy for children with hypophosphatasia. *JCI insight*. 2016;1(9):e85971. doi:10.1172/jci.insight.85971

70. Whyte MP, Rockman-Greenberg C, Ozono K, et al. Asfotase alfa treatment improves survival for perinatal and infantile hypophosphatasia. *J Clin Endocrinol Metab*. 2016;101(1):334-342. doi:10.1210/jc.2015-3462

71. Jones SA, Rojas-Caro S, Quinn AG, et al. Survival in infants treated with sebelipase Alfa for lysosomal acid lipase deficiency: An open-label, multicenter, dose-escalation study. *Orphanet J Rare Dis*. 2017;12(1):25. doi:10.1186/s13023-017-0587-3

72. Ory DS, Ottinger EA, Farhat NY, et al. Intrathecal 2-hydroxypropyl-beta-cyclodextrin decreases neurological disease progression in Niemann-Pick disease, type C1: A non-randomised, open-label, phase 1-2 trial. *Lancet*. 2017;390(10104):1758-1768. doi:10.1016/S0140-6736(17)31465-4

73. Tardieu M, Zerah M, Gougeon M-L, et al. Intracerebral gene therapy in children with mucopolysaccharidosis type IIIB syndrome: An uncontrolled phase 1/2 clinical trial. *Lancet Neurol*. 2017;16(9):712-720. doi:10.1016/S1474-4422(17)30169-2

74. Kinane TB, Mayer OH, Duda PW, Lowes LP, Moody SL, Mendell JR. Long-term pulmonary function in Duchenne muscular dystrophy: Comparison of eteplirsen-treated patients to natural history. *J Neuromuscul Dis*. 2018;5(1):47-58. doi:10.3233/JND-170272

75. Connolly AM, Zaidman CM, Golumbek PT, et al. Twice-weekly glucocorticosteroids in infants and young boys with Duchenne muscular dystrophy. *Muscle Nerve*. 2019;59(6):650-657. doi:10.1002/mus.26441

76. De Vivo DC, Bertini E, Swoboda KJ, et al. Nusinersen initiated in infants during the presymptomatic stage of spinal muscular atrophy: Interim efficacy and safety results from the Phase 2 NURTURE study. *Neuromuscul Disord*. 2019;29(11):842-856. doi:10.1016/j.nmd.2019.09.007

77. Hoffman EP, Schwartz BD, Mengle-Gaw LJ, et al. Vamorolone trial in Duchenne muscular dystrophy shows dose-related improvement of muscle function. *Neurology*. 2019;93(13):e1312-e1323. doi:10.1212/WNL.0000000000008168

78. Polgreen LE, Lund TC, Braunlin E, et al. Clinical trial of laronidase in Hurler syndrome after hematopoietic cell transplantation. *Pediatr Res*. 2020;87(1):104-111. doi:10.1038/s41390-019-0541-2

79. Smith EC, Conklin LS, Hoffman EP, et al. Efficacy and safety of vamorolone in Duchenne muscular dystrophy: An 18-month interim analysis of a non-randomized open-label extension study. *PLoS Med*. 2020;17(9):e1003222. doi:10.1371/journal.pmed.1003222

80. Sondhi D, Kaminsky SM, Hackett NR, et al. Slowing late infantile Batten disease by direct brain parenchymal administration of a rh.10 adeno-associated virus expressing CLN2. *Sci Transl Med*. 2020;12(572):eabb5413. doi:10.1126/scitranslmed.abb5413

81. Wagner KR, Abdel-Hamid HZ, Mah JK, et al. Randomized phase 2 trial and open-label extension of domagrozumab in Duchenne muscular dystrophy. *Neuromuscul Disord*. 2020;30(6):492-502. doi:10.1016/j.nmd.2020.05.002

82. Darras BT, Masson R, Mazurkiewicz-Beldzinska M, et al. Risdiplam-treated infants with type 1 spinal muscular atrophy versus historical controls. *N Engl J Med*. 2021;385(5):427-435. doi:10.1056/NEJMoa2102047

83. Deiva K, Ausseil J, de Bournonville S, et al. Intracerebral gene therapy in four children with Sanfilippo B syndrome: 5.5-year follow-up results. *Hum Gene Ther*. 2021;32(19-20):1251-1259. doi:10.1089/hum.2021.135

84. Gentner B, Tucci F, Galimberti S, et al. Hematopoietic stem- and progenitor-cell gene therapy for Hurler syndrome. *N Engl J Med*. 2021;385(21):1929-1940. doi:10.1056/NEJMoa2106596

85. Ghosh A, Rust S, Langford-Smith K, et al. High dose genistein in Sanfilippo syndrome: A randomised controlled trial. *J Inherit Metab Dis*. 2021;44(5):1248-1262. doi:10.1002/jimd.12407

86. Tsabari R, Simchovitz E, Lavi E, et al. Safety and clinical outcome of tamoxifen in Duchenne muscular dystrophy. *Neuromuscul Disord*. 2021;31(9):803-813. doi:10.1016/j.nmd.2021.05.005

87. Vijay S, Brassier A, Ghosh A, et al. Long-term survival with sebelipase alfa enzyme replacement therapy in infants with rapidly progressive lysosomal acid lipase deficiency: Final results from 2 open-label studies. *Orphanet J Rare Dis*. 2021;16(1):13. doi:10.1186/s13023-020-01577-4

88. Kato M, Kada A, Shiraishi H, et al. Sirolimus for epileptic seizures associated with focal cortical dysplasia type II. *Ann Clin Transl Neurol*. 2022;9(2):181-192. doi:10.1002/acn3.51505

89. Magnani A, Semeraro M, Adam F, et al. Long-term safety and efficacy of lentiviral hematopoietic stem/progenitor cell gene therapy for Wiskott-Aldrich syndrome. *Nat Med*. 2022;28(1):71-80. doi:10.1038/s41591-021-01641-x

90. Strauss KA, Farrar MA, Muntoni F, et al. Onasemnogene abeparvovec for presymptomatic infants with two copies of SMN2 at risk for spinal muscular atrophy type 1: The phase III SPR1NT trial. *Nat Med*. 2022;28(7):1381-1389. doi:10.1038/s41591-022-01866-4

91. Strauss KA, Farrar MA, Muntoni F, et al. Onasemnogene abeparvovec for presymptomatic infants with three copies of SMN2 at risk for spinal muscular atrophy: The phase III SPR1NT trial. *Nat Med*. 2022;28(7):1390-1397. doi:10.1038/s41591-022-01867-3

92. Crawford TO, Swoboda KJ, De Vivo DC, et al. Continued benefit of nusinersen initiated in the presymptomatic stage of spinal muscular atrophy: 5-year update of the NURTURE study. *Muscle Nerve*. 2023;68(2):157-170. doi:10.1002/mus.27853

93. Henzi BC, Schmidt S, Nagy S, et al. Safety and efficacy of tamoxifen in boys with Duchenne muscular dystrophy (TAMDMD): A multicentre, randomised, double-blind, placebo-controlled, phase 3 trial. *Lancet Neurol*. 2023;22(10):890-899. doi:10.1016/S1474-4422(23)00285-5

94. Seo J-H, Kosuga M, Hamazaki T, Shintaku H, Okuyama T. Intracerebroventricular enzyme replacement therapy in patients with neuronopathic mucopolysaccharidosis type II: Final report of 5-year results from a Japanese open-label phase 1/2 study. *Mol Genet Metab*. 2023;140(4):107709. doi:10.1016/j.ymgme.2023.107709

95. Shieh PB, Kuntz NL, Dowling JJ, et al. Safety and efficacy of gene replacement therapy for X-linked myotubular myopathy (ASPIRO): A multinational, open-label, dose-escalation trial. *Lancet Neurol*. 2023;22(12):1125-1139. doi:10.1016/s1474-4422(23)00313-7

96. Chiriboga CA, Bruno C, Duong T, et al. JEWELFISH: 24-month results from an open-label study in non-treatment-naive patients with SMA receiving treatment with risdiplam. *J Neurol*. 2024;271(8):4871-4884. doi:10.1007/s00415-024-12318-z

97. Sondhi D, Kaminsky SM, Rosenberg JB, Rostami MR, Hackett NR, Crystal RG. Twenty-year survival analysis of adeno-associated virus vector serotype 2-mediated gene therapy to the central nervous system for CLN2 disease. *Hum Gene Ther*. 2024;36(1-2):28-35. doi:10.1089/hum.2024.182

98. Canada's Drug Agency. *Methods guide for health technology assessment*. 2025. Accessed June 16, 2025. https://www.cda-amc.ca/sites/default/files/MG%20Methods/MG0030-Quantitative-Methods-Manual_Mar%202025.pdf

99. Haute Autorité de Santé. *Real-world studies for the assessment of medicinal products and medical devices*. 2021. Accessed June 16, 2025. https://www.has-sante.fr/upload/docs/application/pdf/2021-10/real-world_studies_for_the_assessment_of_medicinal_products_and_medical_devices.pdf

100. Haute Autorité de Santé. *Methodology for the clinical development of medical devices*. 2021. Accessed June 16, 2025. https://www.has-sante.fr/upload/docs/application/pdf/2021-09/guide_methodology_for_the_clinical_development_of_md.pdf

101. International Council for Harmonization of Technical Requirements for Pharmaceuticals for Human Use. *E10: Choice of control group and related issues in clinical trials*. 2000. Accessed June 16, 2025. https://database.ich.org/sites/default/files/E10_Guideline.pdf

102. International Council for Harmonization of Technical Requirements for Pharmaceuticals for Human Use. *E8 (R1): General considerations for clinical studies*. 2021. Accessed June 16, 2025. https://database.ich.org/sites/default/files/ICH_E8-R1_Guideline_Step4_2021_1006.pdf

103. National Institute for Health and Care Excellence. *NICE real-world evidence framework*. 2022. Accessed June 16, 2025. https://www.nice.org.uk/corporate/ecd9/resources/nice-realworld-evidence-framework-pdf-1124020816837

104. Hernán MA, Dahabreh IJ, Dickerman BA, Swanson SA. The target trial framework for causal inference from observational data: Why and when is it helpful? *Ann Intern Med*. 2025;178(3):402-407. doi:10.7326/annals-24-01871

105. International Council for Harmonization of Technical Requirements for Pharmaceuticals for Human Use. *E9(R1): Addendum on estimands and sensitivity analysis in clinical trials to the guideline on statistical principles for clinical trials*. 2019. Accessed May 1, 2026. https://database.ich.org/sites/default/files/E9-R1_Step4_Guideline_2019_1203.pdf

106. Liu J, Yao M, Wang M, et al. Design, conduct, and analysis of externally controlled trials. *JAMA Netw Open*. 2025;8(9):e2530277-e2530277. doi:10.1001/jamanetworkopen.2025.30277

107. Zayadi A, Edge R, Parker CE, et al. Use of external control arms in immune-mediated inflammatory diseases: A systematic review. *BMJ Open*. Dec 9 2023;13(12):e076677. doi:10.1136/bmjopen-2023-076677

108. Hernán MA, Wang W, Leaf DE. Target Trial Emulation: A framework for causal inference from observational data. *JAMA*. 2022;328(24):2446-2447. doi:10.1001/jama.2022.21383

109. Hernán MA, Robins JM. Using big data to emulate a target trial when a randomized trial is not available. *Am J Epidemiol*. 2016;183(8):758-764. doi:10.1093/aje/kwv254

110. Hogervorst MA, Soman KV, Gardarsdottir H, Goettsch WG, Bloem LT. Analytical methods for comparing uncontrolled trials with external controls from real-world data: A systematic literature review and comparison with European regulatory and health technology assessment practice. *Value Health*. 2025;28(1):161-174. doi:10.1016/j.jval.2024.08.002

Table 1. Documents included in extraction and number of statements extracted

Document with guidance	Number of extracts
Canada's Drug Agency	
Methods guide for health technology assessment (2025)[98]	1
French National Authority for Health	
Real-world studies for the assessment of medicinal products and medical devices (2021)[99]	1
Methodology for the clinical development of medical devices (2021)[100]	7
International Council for Harmonization of Technical Requirements for Pharmaceuticals for Human Use	
Guideline E10: Choice of control group and related issues in clinical trial (2000)[101]	11
Guideline E8 (R1): General considerations for clinical studies (2025)[102]	2
National Institutes for Clinical Excellence	
NICE real-world evidence framework (2022)[103]	8
United States Food and Drug Administration	
Considerations for the design and conduct of externally controlled trials for drug and biological products (2023)[20]	40

Table 2. Guidance statements regarding criteria for using an externally controlled trial

ID	Topic	Synthesized guidance statement	Verbatim extract IDs[a]
1	Possibility of randomization	An externally controlled trial is acceptable in cases where a randomized comparison is not feasible or ethical or when direct comparative evidence is unavailable.	2A, 4A, 4C, 4G, 7I
2	Disease	Disease progression should be well understood and highly predictable to incorporate an external control.	1E, 6A, 7A, 7I
3	Disease, Fit for purpose data	Prognostic factors should be well understood and available in the external control data source.	1E, 1F, 7J
4	Effect Size	An external control is most appropriate when the anticipated treatment effect is large.	1AJ, 7A, 7J
5	Outcomes	Endpoints should be objective and the impact of baseline and treatment variables on the endpoint well characterized.	7B, 7J
6	Comparability across arms	If applicable, treatment attributes and treatment-related factors should be assessed for comparability across the internal and external control arms prior to the conduct of an externally controlled trial.	1W, 1X

[a]See Table SI 1 for corresponding verbatim extracts.

Table 3. Guidance statements mapping to ROBINS-I V2 domain: Risk of bias due to confounding

ID	Topic	Synthesized guidance statement	Verbatim extract IDs[a]	Comparison to externally controlled trials in pediatric genetic rare disease
7	Fit for purpose data, Comparability across arms	Where possible, select external control data with a source population similar to the internal arm to minimize differences in key study variables between the internal and external arms, including patient characteristics, diagnostic standards, data collection, healthcare, geographic regions and time periods. Consider the impact of differences between arms across these factors on result interpretation.	1G, 1S, 1T,1U, 4E, 5A, 5D, 5F, 6A, 7C, 7J	Difficult to ascertain from published articles.
8	Fit for purpose data	Where possible, select external control data with detailed individual patient data on baseline demographics/status, concomitant therapy and course on study.	7C	All 18 articles selected external control data with individual-level patient data.[b]
9	Comparability across arms	The same inclusion criteria should be applied to the internal and external arms.	1I, 5E	All 18 articles reported applying the same or highly similar inclusion criteria to their internal and external arms.
10	Comparability across arms	Prognostic factors should be compared and shown to be sufficiently similar across the internal and external arms to permit an unbiased assessment.	1V	15 of 18 articles included a comparison of baseline covariates between the internal and external arms. Of the 15, 4 made a comparison prior to applying a specific method to address imbalances, 9 made a comparison after applying a specific method to address imbalances and 2 made a comparison both before and after.
11	Comparability across arms, Analytic method	The analytic methods chosen should identify and manage sources of confounding and bias in both the internal and external arms, including differences in baseline characteristics or confounding variables between the internal and external arms.	1H, 1J, 1AG, 4D, 4F	All 18 articles applied a specific analytic method to address differences in baseline covariates, beyond adjusting for age or baseline outcome score.[b]
12	Comparability across arms, Analytic method	In addition to the implementation of analytic methods to balance trial arm population, additional analyses should evaluate comparability between arms on important variables.	1AI	3 of 18 articles discussed the success of their analytic method chosen to address differences in baseline covariates.
13	Additional analyses	Quantitative or qualitative bias analyses should be planned a priori and performed to assess impact of confounding factors and sources of bias.	1AK, 5I	No articles reported a bias analysis.

[a]See Table SI 1 for corresponding verbatim extracts.

[b]A condition of inclusion for the prioritized subset of articles.

Table 4. Guidance statements mapping to ROBINS-I V2 domain: Risk of bias due to outcome measurement

ID	Topic	Synthesized guidance statement	Verbatim extract IDs[a]	Comparison to externally controlled trials in pediatric genetic rare disease
14	Outcomes	Selecting objective and reliable measurements for the data of interest is the ideal strategy to avoid misclassification.	1AP	Difficult to ascertain from published articles.
15	Protocol/Statistical analysis plan	Timing and duration of outcome assessments should be established prior to the selection of external data.	1P	Difficult to ascertain from published articles.
16	Comparability across arms, Outcomes	Timing and method of outcome assessment should be consistent across the internal and external arms.	1N, 1O, 1P, 1AA, 7D	All 18 articles reported using the same outcome measurement instrument across the internal and external arms. 8 of 18 articles reported the same assessment time points between the internal and external arms for all outcomes.
17	Protocol/Statistical analysis plan	The study design and statistical analysis plan should be blinded to any observed external control data, except planned feasibility analyses.	1AE	Difficult to ascertain from published articles.
18	Outcomes	Where possible, consider blinding for outcome assessment in the internal and external arms.	1M, 7H	1 of 18 studies reported blinding in the assessment of at least one outcome.

[a]See Table SI 1 for corresponding verbatim extracts.

Table 5. Guidance statements mapping to ROBINS-I V2 domain: Risk of bias due to the selection of participants or missing data

ID	Topic	Synthesized guidance statement	Verbatim extract IDs[a]	Comparison to externally controlled trials in pediatric genetic rare disease
	Selection of participants			
19	Comparability across arms, Outcomes	Immortal time bias should be considered when establishing an index date and follow-up period for outcome assessment across internal and external arms - both should be consistent and comparable across the internal and external arms.	1K, 1L, 1Y	12 of 18 articles did not specify how the index date of their external arm(s) was selected.
	Missing data			
20	Missing data	The extent of and reason for missing data in external control data and their potential impact on result interpretation should be captured and considered a priori.	1Q, 1AB, 1AM	Difficult to ascertain from published articles.
21	Analytic method, Missing data	The analytic method should include a strategy to address missing data in external arm(s).	1AL	1 of 18 studies specified how missing data in the external arm would be addressed.
22	Missing data	The extent of missing data in both the internal and external arms should be assessed to examine their impact on study results.	1Q, 1AB, 1AM	4 of 18 articles reported on missingness in baseline covariates of their external arm(s). 4 of 18 studies reported on missingness in baseline covariates of their internal arm(s). 10 of 18 articles reported on missingness in the outcome measurement of their external arm(s). 12 of 18 articles reported on missingness in the outcome measurement of their internal arm(s).

[a]See Table SI 1 for corresponding verbatim extracts.

Table 6. Guidance statements mapping to multiple ROBINS-I V2 domains

ID	Topic	Synthesized guidance statement	Verbatim extract IDs[a]	Comparison to externally controlled trials in pediatric genetic rare disease
23	Protocol/Statistical analysis plan	External control selection, design elements (such as baseline eligibility, exposure definition, endpoints) and analyses (such primary and secondary endpoints, power calculations, sample size, approaches to minimizing missing data and plans to control the chance of erroneous conclusions) should be planned and specified prior to the conduct of an externally controlled trial.	1A, 1B, 1D, 1AC, 3A, 4B, 4D, 4F, 6B, 7E, 7F	14 of 18 articles pre-specified the inclusion of an external control. 12 of 13 articles with an accessible protocol pre-specified their external control, design elements and analysis prior to the conduct of their externally controlled trial.
24	Protocol/Statistical analysis plan	Avoid revisions to the statistical analysis plan where possible. If changes are implemented, document, report and justify them.	1AF	Difficult to ascertain from published articles.
25	Protocol/Statistical analysis plan	Target trial emulation should be considered, if unfeasible, trade-offs should be discussed.	5D	No articles reported considering or using target trial emulation.
26	Protocol/Statistical analysis plan	Consider using the estimand framework to help design an externally controlled trial and account for any intercurrent events.	1C, 1AO	No articles reported using the estimand framework.
27	Fit for purpose data	Where possible, consider the use of data from a previous clinical trial to create an external control arm.	1R, 5C	2 of 18 articles reported using data from a previously conducted clinical trial to create an external control arm.
28	Pooling data sources	Multiple controls can be studies when no single optimal external control exits, if pre-planned and analyzed/interpreted conservatively.	7G	5 of 18 articles included two or more external control arms in their analysis.
29	Pooling data sources	If multiple potential sources of external data are suitable, pooling should only be done in cases where there is limited heterogeneity in terms of coverage and data quality.	5G	3 of 18 articles reporting pooling data to create an external control.
30	Comparability across arms	Differences in variable definitions and data collection across the internal and external arms should be reported.	5J	All 18 articles reported using the same outcome measurement instrument across the internal and external arms.
31	Analytic method	Consider a conservative approach, including the potential for a more stringent statistical significance level, given the potential for bias in favour of the treatment.	7K	Difficult to ascertain from published articles.
32	Analytic method	Justify the analytic method selected for comparison with the external control (strengths and limitations) and examine its assumptions.	1AH, 1AG	Difficult to ascertain from published articles.
33	Additional analyses	Sensitivity analyses should be used to evaluate the potential impact of confounders, missing data and other potential sources of bias.	1AN, 4F	5 of 18 articles conducted a sensitivity analysis, each of which were used to evaluate the potential impact of confounders, missing data or other potential biases.

[a]See Table SI 1 for corresponding verbatim extracts

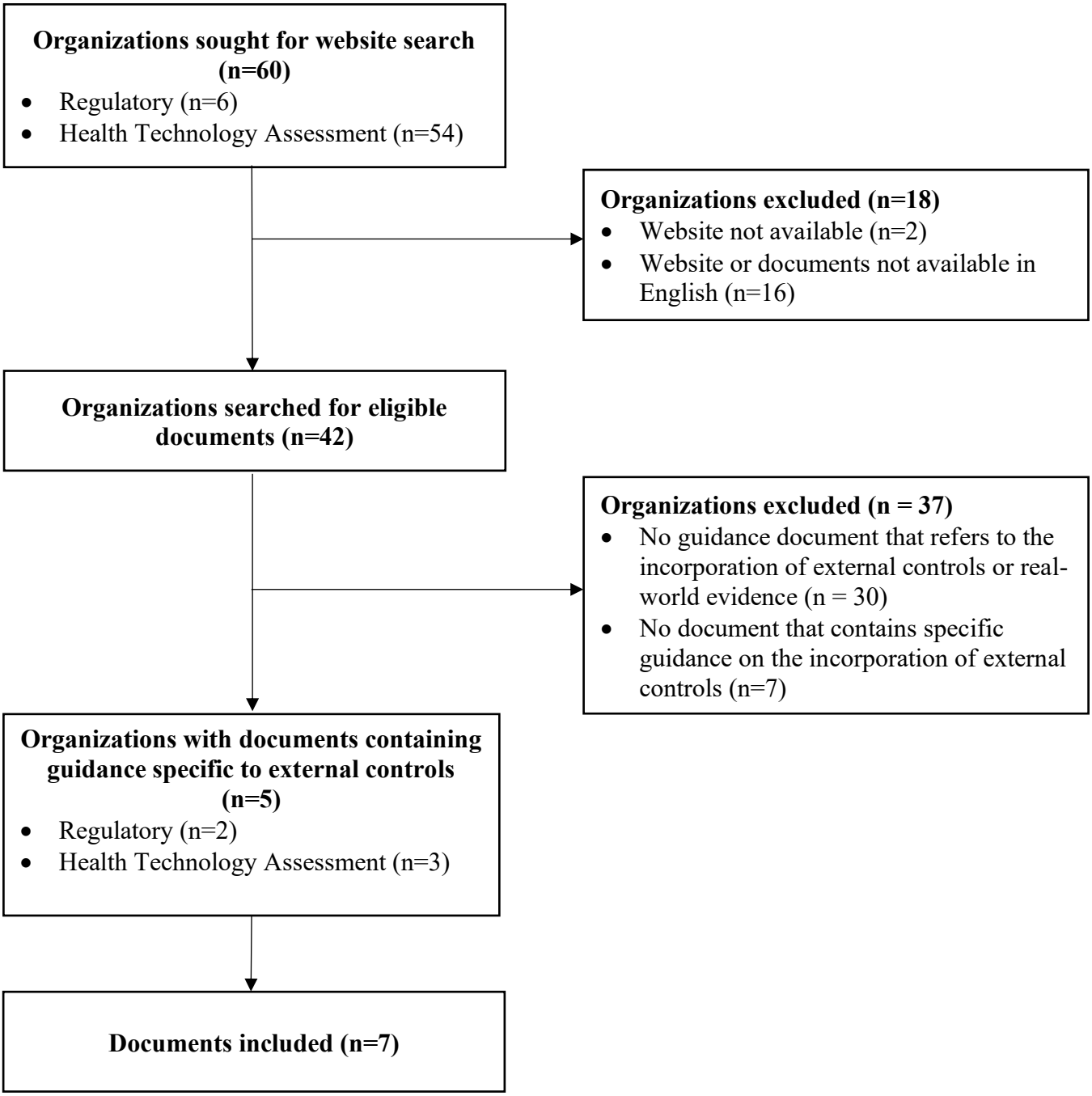


Figure 1. PRISMA flow diagram depicting the identification, screening and selection of documents included in the focused literature search.

Supplementary Methods 1. A complete list of eligible regulatory bodies and HTA organizations

Organization
Regulatory - ICH
International Council for Harmonization of Technical Requirements for Pharmaceuticals for Human Use
Regulatory – ICH founding regulatory member
European Medicines Agency
(United States) Food and Drug Administration
(Japan) Ministry of Health and Labour
Regulatory – ICH standing regulatory member
Health Canada
Swissmedic
HTA – Member of INAHTA
Agency for Care Effectiveness
Adelaide Health Technology Assessment
Agency for Health Technology Assessment and Tariff System
Agency for Healthcare Research and Quality
Agency of Health Quality and Assessment of Catalonia
Andalusian Health Technology Assessment Department
Australian Safety and Efficacy Register of New Interventional Procedures - Surgical
Austrian Institute for Health Technology Assessment
Austrian National Public Health Institute
Basque Office for Health Technology Assessment
Belgian Health Care Knowledge Centre
Canada's Drug Agency
Center for Drug Evaluation
Center for Outcomes Research and Economic Evaluation for Health
Colombian agency of health technology assessment
Committee for Evaluation and Dissemination of Innovative Technologies – Paris University Hospital
Danish Healthcare Quality Institute
Emilia-Romagna Regional Health Authority
Federal Joint Committee
Federal Office of Public Health
Finnish Coordinating Center for Health Technology Assessment
French National Authority for Health
Galician Agency for Health Technology Assessment
General Directorate of Medicines, Supplies and Drugs
Health Assessment Division of the Ministry of Health
Health Information and Quality Authority
Health technology Assessment and Innovation unit - A.Gemelli Teaching Hospital
Health Technology Assessment of India
Health Technology Wales
Healthcare Improvement Scotland
Institute for Clinical Effectiveness and Health Policy
Institute for Health Sciences of Aragon
Institute for Quality and Efficiency in Health Care

Institute of Health Economics
Italian National Agency for Regional Healthcare Services
Malaysian Health Technology Assessment Section
National Authority for Assessment and Accreditation in Healthcare
National Committee for Technology Incorporation
National Evidence-based Healthcare Collaborating Agency
National Health Care Institute
National Institute for Clinical Excellence
National Institute for Health and Care Research
National Institute for Value and Technologies in Healthcare
National Institute of Excellence in Health and Social Services
National Supplementary Health Agency
Netherlands Organisation for Health Research and Development
Norwegian Institute of Public Health
Ontario Health
Pharmac
Republican Center for Health Development
Spanish Network of Agencies for Assessing National Health System Technologies and Performance
Swedish Agency for Health Technology Assessment and Assessment of Social Services

Note. HTA: Health Technology Assessment, ICH: International Council for Harmonization of Technical Requirements for Pharmaceuticals for Human Use, INAHTA: International Network of Agencies for Health Technology Assessment.

Supplementary Methods 2. A complete list of keywords used in website search

External control

External comparison

External comparator

Historical control

Historical comparison

Historical comparator

Natural history control

Natural history comparison

Natural history comparator

Synthetic control

Synthetic comparison

Synthetic comparator

Real world data

Real world evidence

Table SI 1. Verbatim guidance statements extracted from included documents and assigned ID

ID	Verbatim statement	Page number
Considerations for the design and conduct of externally controlled trials for drug and biological products[20]		
1A	"Sponsors should finalize a study protocol before initiating the externally controlled trial, including selection of the external control arm and analytic approach, rather than selecting an external control arm after the completion of a single-arm trial."	4
1B	"Specific design elements to prespecify in the protocol (i.e., before conducting an externally controlled trial) include suitable study data sources, baseline eligibility (inclusion and exclusion) criteria,19 appropriate exposure definitions and windows, well-defined and clinically meaningful endpoints, cogent analytic plans, and approaches to minimize missing data and sources of bias."	4
1C	"The estimand framework—involving a precise description of the treatment effect reflecting the clinical question posed by the study objective—can be used to help design an externally controlled trial."	5
1D	"A specific design consideration for externally controlled trials involves prespecifying plans regarding how to measure and analyze data on important confounding factors and sources of bias."	5
1E	"Conceptually, when seeking to provide evidence of effectiveness using an externally controlled trial design, a thorough understanding is needed—but is often difficult to verify—regarding the natural history of the disease involved and relevant prognostic factors influencing outcomes."	5
1F	"Accordingly, before deciding whether an externally controlled trial is a suitable design to answer the research question of interest, sponsors should confirm that recognized, important prognostic characteristics can be assessed in the data sources that will be used in an externally controlled trial."	5
1G	"Specifically, the source population for the external control arm should be as comparable as possible to the treatment arm population, given that controlling for differences between the two study arms (see section III.C) becomes more challenging with increasingly dissimilar populations."	5
1H	"Although unmeasured confounding, lack of blinding, and other sources of bias cannot be eliminated in externally controlled trials, an assessment of the extent of confounding and bias, along with analytic methods to reduce the impact of such bias, are critically important in the conduct of such trials."	5
1I	"Accordingly, the protocol for an externally controlled trial should include specific plans for evaluating eligibility criteria to determine if the criteria can be applied in a manner that allows for selection of similar patients in the treatment and external control groups, recognizing the limitations of information available in many RWD sources."	6
1J	"These and other health care delivery factors—at the level of the patient, provider, or health system—can influence treatment selection. Such factors should be identified and accounted for adequately in externally controlled trials; otherwise, a different design approach (e.g., randomized controlled trial) should be considered."	7

1K	"Determination of the index date in the treatment arm and the external control arm should avoid analyses that include a period of time (immortal time) during which the outcome of interest could not have occurred in one of the two arms."	8
1L	"When assessing bias that may be introduced related to immortal time in an externally controlled trial, the clinical circumstances related to assigning the index date should be considered."	8
1M	"Accordingly, whenever possible and for suitable endpoints, the outcome should be assessed blinded to treatment status."	8
1N	"Sponsors should seek to assess outcomes consistently across the treatment arm and the external control arm for the results of an externally controlled trial to be credible."	8
1O	"When considering outcomes in externally controlled trials, sponsors should also evaluate the consistency of timing of outcome assessments in the treatment arm compared to the external control arm."	9
1P	"Accordingly, sponsors should first establish for what total duration of time and at what intervals the outcome of interest should be assessed in the analysis of data from an externally controlled trial. Based on such determinations, sponsors can then evaluate whether the availability and timing of outcome assessments are sufficient and comparable across both arms of the externally controlled trial for the research hypothesis being tested."	9
1Q	"Furthermore, when using data from other clinical trials as an external control arm, sponsors should consider the extent of and reason for any missing data and how the interpretability of study results may be affected."	10
1R	"Accordingly, sponsors should assess whether use of data from a specific clinical trial is justified as an external control arm when planning an externally controlled trial."	11
1S	"It is important to consider whether and how different time frames in the treatment arm and the external control arm impact the interpretability of study findings."	12
1T	"A balance of participants or patients across geographic regions and health care systems in an externally controlled trial, when possible, can help reduce the impact of confounding based on such differences."	12
1U	"Sponsors should consider the diagnostic standards used and whether relevant clinical tests to establish a diagnosis were conducted and reported equally across the compared arms."	12
1V	"Based on demographic and clinical characteristics—and if sufficient knowledge of relevant prognostic factors is available—prognostic indicators for the participants or patients in each arm of the trial should be evaluated and shown to be of sufficient similarity to permit an unbiased assessment of the treatment-outcome association."	12
1W	"Attributes of the treatment of interest—including drug formulation, dose, route of administration, timing, frequency, and duration as well as specific rules for dose modifications, interruptions, discontinuations, and adherence—will have been prespecified or measured in the treatment arm. In contrast, specific aspects of a comparator treatment (as applicable) in the external control arm may not have been protocol-driven depending on the data source. Accordingly, sponsors should assess whether the external control arm data can be meaningfully compared to the treatment arm data."	12
1X	"Various treatment-related considerations, when relevant, include (1) previous treatments received (e.g., lines of therapy in patients with cancer), (2) medications received concomitantly that can affect the outcome of interest, or (3) predictive biomarkers (e.g.,	12

	genomic testing) related to the treatment of interest. When differentially distributed across groups being compared, such factors can threaten an assessment of the drug-outcome association."	
1Y	"Designation of the index date should be consistent between the treatment arm and the external control arm, and the duration of follow-up periods should be comparable across compared arms."	12
1AA	"Sponsors should be able to apply the same criteria for the evaluation and timing of outcome assessments across both arms of the externally controlled trial."	13
1AB	"The extent of missing data in the external control arm should be assessed before conducting an externally controlled trial to evaluate feasibility (when such data are available). When analyzing results from such a trial, the extent of missing data in both the treatment and external control arms should be assessed to examine the potential impact of missing data."	
1AC	"Before conducting an externally controlled trial, sponsors should develop a statistical analysis plan that prespecifies analyses of interest, such as analyses of primary and secondary endpoints, calculations of statistical power and sample size, and plans to control the chance of erroneous conclusions (e.g., to control the overall type I error probability)."	13
1AE	"In addition, decisions regarding the study design and statistical analysis plan for an externally controlled trial should be blinded to any observed external control data (e.g., from an existing RWD source), with the exception of planned feasibility analyses, such as evaluating the availability of key variables or missing data."	13
1AF	"During the conduct of an externally controlled trial, and specifically when analyzing data already collected, changes to the statistical analysis plan are discouraged. If such changes are nonetheless implemented, any revisions should be date-stamped and the corresponding rationale provided and discussed with the relevant FDA review division."	13
1AG	"Sponsors should provide a justification for the analytic methods selected as well as a description of the strengths and limitations of the methods used to assess the effect of treatment. In general, the analytic method used should identify and manage sources of confounding and bias, including a strategy to account for differences in baseline factors and confounding variables between trial arms."	13
1AH	"Various statistical methodologies may be appropriate for these types of comparisons, each with a corresponding level of complexity regarding approaches to account for bias. The assumptions involved should be made explicit, and sensitivity analyses as well as model diagnostics should be conducted to examine such assumptions."	13-14
1AI	"Even when employing analytic methods to balance the trial arm populations, sponsors should propose additional analyses to evaluate the actual comparability between the external control and treatment arms for important covariates. Determining similarity across trial arms will require selection of specific population characteristics to compare, a method for the comparison, and criteria to demonstrate similarity."	14
1AJ	"Consideration should also be given, based on available scientific data, to the anticipated effect size for analyses of the primary endpoint. Especially when the anticipated effect size is modest, an externally controlled trial may not be an appropriate study design because of concerns for bias affecting the results."	14

1AK	"In addition, sponsors should develop a priori plans for assessing the impact of confounding factors and sources of bias, with quantitative or qualitative bias analyses used to evaluate these concerns."	14
1AL	"The proposed analytical methods should include a strategy for dealing with missing data, including data that may not be available in a chosen data source based on the type and frequency of assessments conducted during the patient encounter, patients no longer being followed, or other reasons."	14
1AM	"To understand the potential impact of missing data, externally controlled trials should be designed to capture and analyze information relevant to the missing data (e.g., available characteristics of patients with and without missing data)."	15
1AN	"In addition, sensitivity analyses should be used to evaluate the potential impact of plausible violations in missing data assumptions on the results of the primary and other key analyses."	15
1AO	"The study analysis plan and an appropriate estimand should account for any intercurrent event that can be considered potentially related to both the treatment and outcome of interest, recognizing that certain intercurrent events may be difficult to detect in external control datasets."	15
1AP	"Although analytical modeling methods could be used to assess the potential impact of misclassification, the best strategy to avoid bias is to use objective and reliable measurements for the data of interest."	15
Methods guide for health technology assessment[98]		
2A	"Observational and RWE studies, including single-arm clinical trials in which there is an external control arm (historical control or contemporaneous cohort), may be considered on a case-by-case basis when it is not feasible or ethical to conduct an RCT, or when there are uncertainties in the clinical interventional trial and/or indirect comparison evidence."	13
Real-world studies for the assessment of medicinal products and medical devices[99]		
3A	"In addition, real-world data may be generated or collected to serve as an external control group for a non-randomised clinical trial. In this case, the implementation of an external comparison must be anticipated and scheduled in advance in order to improve its robustness and ensure it is part of a deductive reasoning approach."	10
Methodology for the clinical development of medical devices[100]		
4A	"External comparative studies should be reserved for situations where randomisation is not feasible. In the absence of direct comparison, an indirect comparison, conducted on the basis of defined and validated methodological principles, may be taken into account."	52
4B	"The external comparison must be formalised in the study protocol with prior choice of the comparison reference as well as the formal comparison method: test against a standard, adjustment with individual data, synthetic control group or Matching-Adjusted Indirect Comparisons (MAIC)."	52
4C	"Single-arm studies with external comparisons can be used when the target population is small and inclusion is difficult."	52
4D	"To be acceptable, these studies must be able to guarantee the absence of residual confounding factors. This comparison is limited by 1) the post hoc choice of the reference for the comparison, 2) the confounding bias, which necessarily requires an adjustment approach, and 3) other biases: selection, measurement and missing data among others."	52

4E	"For the control arm where retrospective data collection is implemented, the following limitations should be taken into account: - modification over time of the therapeutic strategy, changing the disease history; - increase in operator and patient experience over time; - different patient characteristics at inclusion between a historical control group and a contemporary group."	52
4F	"In order to limit bias, it is necessary to define the external control in advance and justify it, i.e. to ensure an unbiased choice of the reference value of the external control (based on a systematic review of reference studies, taking into account the statistical uncertainty of these reference values and retaining the most unfavourable value), identify all potential confounders (prognostic factors of the endpoints and modifiers of the treatment effect) by systematic review, take into account these confounders and potential biases, and perform sensitivity analyses."	53
4G	"Data from real-world studies can also be used, under certain conditions, as a control arm in an indirect comparison study, particularly where there is a registry or systematic data collection. However, this use is difficult when the target patients for a new MD differ substantially from those of the reference therapeutic strategies. In the field of MDs, single-arm studies can enable, for example, the assessment of the evolution of a device, the interest of which was initially demonstrated by a randomised clinical trial, or when the number of patients available is too small to carry out randomised clinical trials."	53
NICE real-world evidence framework[103]		
5A	"For studies using external control, select and curate data to minimise differences between data sources including availability and operational definitions of key study variables, data collection processes, patient characteristics, treatment settings, care pathways, and time periods, and consider the implications for study quality and relevance."	58
5C	"External controls can also be formed from data from previous clinical trials. A potential advantage of such studies is greater similarity in patient inclusion criteria, follow up and outcome determination. Often only aggregate rather than individual patient-level data will be available from previous trials."	62
5D	"Studies should aim to emulate the target trial as closely as possible and, if this is not possible, trade-offs should be clearly described. In some cases, a data source may not be of sufficient relevance or quality to allow trial emulation. This can be particularly problematic for studies using real-world data to form an external control because differences in terms of patients, settings, care, data collection and time periods can limit the comparability between the trial and the real-world data."	65
5E	"For most studies, the eligibility criteria should mimic a hypothetical pragmatic trial by reflecting the clinical pathways (including diagnostic tests) and patients seen in routine care in the NHS. For external control studies, the focus should be on matching the eligibility criteria from the interventional study rather than the broader target population."	65
5F	"Data on comparators would ideally come from the same period as the intervention as well as from the same healthcare system and settings. This is to minimise any differences between treatment groups resulting from differences in care access, pathways (including diagnostic tests) or time-based trends in outcomes."	66
5G	"If there are multiple potential sources of suitable real-world data to provide external control to trial data, developers should consider whether to estimate effects separately for each data source or to increase power by pooling data sources. Data sources	74

	should only be pooled when there is limited heterogeneity between sources in terms of coverage and data quality. Individual estimates of effects for each data source should always be provided."	
5I	"Bias analysis may be particularly valuable in studies using real-world data external controls if differences between data collection, settings and time may reduce comparability of data."	79
5J	"For external control studies, differences in variable definitions and data collection should be clearly described."	83
E8 (R1): General considerations for clinical studies[102]		
6A	"Important limitations of the use of external controls are discussed in ICH E10. Particular care is needed to minimise the likelihood of erroneous inference. The use of an external control requires that the disease course is well known and predictable. External control individuals may differ from study participants with respect to demographic and background characteristics (e.g., medical history, concurrent diseases). In addition, external control individuals may differ from participants in the study with respect to concurrent care and the measurement of study outcomes and other data elements. Because the use of internal controls generally mitigates the potential for bias better than external controls, particularly in conjunction with randomisation, the suitability of the use and choice of external control should be carefully considered and justified."	16
6B	"For example, for a single-arm interventional study with an external control, the specifics of the external control should be defined prior to the conduct of the interventional aspect of the study. Pre-specification of the analysis should be in place so that any review of the existing data sources prior to the design of the study does not threaten the study integrity."	18
E10: Choice of control group and related issues in clinical trials[101]		
7A	"A consequence of the recognized inability to control bias is that the potential persuasiveness of findings from externally controlled trials depends on obtaining much more extreme levels of statistical significance and much larger estimated differences between treatments than would be considered necessary in concurrently controlled trials. The inability to control bias restricts use of the external control design to situations in which the effect of treatment is dramatic and the usual course of the disease highly predictable."	25
7B	"In addition, use of external controls should be limited to cases in which the endpoints are objective and the impact of baseline and treatment variables on the endpoint is well characterized."	25
7C	"A control group should be chosen for which there is detailed information, including, where pertinent, individual patient data regarding demographics, baseline status, concomitant therapy, and course on study. The control patients should be as similar as possible to the population expected to receive the test drug in the study and should have been treated in a similar setting and in a similar manner, except with respect to the study therapy."	25
7D	"Study observations should use timing and methodology similar to those used in the control patients."	25
7E	"To reduce selection bias, selection of the control group should be made before performing comparative analyses; this may not always be feasible, as outcomes from these control groups may have been published."	25
7F	"Any matching on selection criteria or adjustments made to account for population differences should be specified prior to selection of the control and performance of the study."	25

7G	"Where no obvious single optimal external control exists, it may be advisable to study multiple external controls, providing that the analytic plan specifies conservatively how each will be used in drawing inferences (e.g., study group should be substantially superior to the most favorable control to conclude efficacy)."	25
7H	"In some cases, it may be useful to have an independent set of reviewers reassess endpoints in the control group and in the test group in a blinded manner according to common criteria."	25
7I	"An externally controlled trial should generally be considered only when prior belief in the superiority of the test therapy to all available alternatives is so strong that alternative designs appear unacceptable and the disease or condition to be treated has a well-documented, highly predictable course. It is often possible, even in these cases, to use alternative, randomized, concurrently controlled designs."	26
7J	"Externally controlled trials are most likely to be persuasive when the study endpoint is objective, when the outcome on treatment is markedly different from that of the external control and a high level of statistical significance for the treatment-control comparison is attained, when the covariates influencing outcome of the disease are well characterized, and when the control closely resembles the study group in all known relevant baseline, treatment (other than study drug), and observational variables."	26
7K	"However, despite the use of a single treatment group in an externally controlled trial, the estimate of the external control group outcome always should be made conservatively, possibly leading to a larger sample size than would be needed in a placebo-controlled trial. Great caution (e.g., applying a more stringent significance level) is called for because there are likely to be both identified and unidentified or unmeasurable differences between the treatment and control groups, often favoring treatment."	27

Table SI 2. Coding of verbatim guidance statements by assigned ID

ID	Category	Topic(s)	Message(s)
Considerations for the design and conduct of externally controlled trials for drug and biological products[20]			
1A	Design consideration	Protocol/Statistical analysis plan	External control selection and analysis should be planned and specified a priori.
1B	Design consideration	Protocol/Statistical analysis plan	Prespecify external control data source and design elements in protocol (baseline eligibility, exposure definition, endpoints, approaches to minimizing missing data and other sources of bias, and analytic approach).
1C	Design consideration	Protocol/Statistical analysis plan	Consider using the estimand framework to help design an externally controlled trial.
1D	Design consideration	Protocol/Statistical analysis plan	Analytic method to identify and manage sources of confounding and bias should be identified a priori.
1E	Criteria for using an external control	Disease	Disease progression should be well known and highly predictable to incorporate an external control.
1F	Criteria for using an external control	Fit for purpose data	Prognostic factors should be available in external control data source.
1G	Design consideration	Comparability across arms, Fit for purpose data	Source population for external control arm(s) should be comparable to internal treatment arm(s).
1H	Analysis consideration	Analytic method, Comparability across arms	The analytic methods chosen should identify, assess and manage all sources of confounding and bias in both the internal and external arms.
1I	Design consideration	Comparability across arms	The same inclusion criteria should be applied to the internal and external arms.
1J	Analysis consideration	Comparability across arms	Confounding factors affecting treatment selection should be identified and addressed.
1K	Design consideration	Outcomes	Consider immortal time bias when establishing index date and follow-up period for outcome assessment across the internal and external arms.
1L	Design consideration	Outcomes	Ensure immortal time bias is avoided in establishing the index date for outcome assessment.
1M	Design consideration	Outcomes	Consider blinding for outcome assessment in the internal and external arms, where possible.
1N	Design consideration	Comparability across arms, Outcomes	Method of outcome assessment should be consistent across the internal and external arms.
1O	Design consideration	Comparability across arms, Outcomes	Timing of outcome assessment should be consistent across the internal and external arms.

1P	Design consideration	Comparability across arms, Protocol/Statistical analysis plan, Outcomes	Timing and duration of outcome assessments should be established prior to the selection of external data. Timing and method of outcome assessment should be consistent across the internal and external arms.
1Q	Design consideration	Missing data	The extent and reason for missing data in external arm should be considered.
1R	Design consideration	Fit for purpose data	Where possible, consider the use of data from a previous clinical trial to create an external control arm.
1S	Design consideration	Comparability across arms	Consider how time frame differences and comparability between the internal and external arms may affect results.
1T	Design consideration	Comparability across arms	To reduce bias related to geographic region, participants in the internal and external arms should be balanced across geographic regions, when possible.
1U	Design consideration	Comparability across arms	Consider comparability of diagnostic standards, including clinical test use and reporting across the internal and external arms.
1V	Analysis consideration	Comparability across arms	Prognostic factors should be compared and shown to be sufficiently similar across the internal and external arms to permit an unbiased assessment.
1W	Criteria for using an external control	Comparability across arms	In cases where the external control group received a treatment, attributes of the treatment (drug formulation, dose, administration route, timing, frequency and duration) should be assessed before proceeding.
1X	Criteria for using an external control	Comparability across arms	Distribution of previous treatments received, medications received concomitantly, predictive biomarkers should be similar across the internal and external arms.
1Y	Design consideration	Comparability across arms, Outcomes	Designation of index date and follow up duration across arms should be comparable across the internal and external arms.
1AA	Design consideration	Comparability across arms, Outcomes	Timing and method of outcome assessment should be consistent across the internal and external arms.
1AB	Design considerations, Analysis considerations	Missing data	The extent of and reason for missing data in external control data and their potential impact on result interpretation should be considered a priori. The extent of missing data in both the internal and external arms should be assessed to examine their impact on study results.
1AC	Analysis considerations	Protocol/Statistical analysis plan	Prespecify analyses of interest in the statistical analysis plan (primary and secondary endpoints, power calculations, sample size and plans to control the chance of erroneous conclusions).
1AE	Analysis considerations	Protocol/Statistical analysis plan	The study design and statistical analysis plan should be blinded to any observed external control data, except planned feasibility analyses.
1AF	Analysis considerations	Protocol/Statistical analysis plan	Avoid revisions to the statistical analysis plan where possible. If changes are implemented, document, report and justify them.

1AG	Analysis considerations	Analytic method, Comparability across arms	Analytic method should account for differences in baseline and confounding variables between internal and external arms. Justify analytic method selected for comparison with external control and describe its strengths and limitations.
1AH	Analysis considerations	Analytic method	Identify and examine assumptions involved in analytic method chosen for comparison with external control.
1AI	Analysis considerations	Analytic method, Comparability across arms	In addition to the implementation of analytic methods to balance trial arm population, additional analyses should evaluate comparability between arms on important variables.
1AJ	Criteria for using an external control	Effect size	An external control is most appropriate when the anticipated treatment effect is large.
1AK	Analysis considerations	Additional analyses	Quantitative or qualitative bias analyses should be planned a priori and performed to assess impact of confounding factors/sources of bias.
1AL	Analysis considerations	Analytic method, Missing data	The analytic method should include a strategy to deal with missing data in external arm(s).
1AM	Design consideration	Missing data	The extent of and reason for missing data in external control data and their potential impact on result interpretation should be captured and considered a priori.
1AN	Analysis considerations	Additional analyses	Sensitivity analyses should be used to evaluate the potential impact of missing data.
1AO	Design consideration	Protocol/Statistical analysis plan	The statistical analysis plan and estimand should account for any intercurrent event (part of estimand framework) potentially related to treatment and outcome.
1AP	Design consideration	Outcomes	Selecting objective and reliable measurements for the data of interest is the ideal strategy to avoid misclassification.
Methods guide for health technology assessment[98]			
2A	Criteria for using an external control	Possibility of randomization	An externally controlled trial is acceptable in cases where a randomized comparison is not feasible or ethical or when direct comparative evidence is unavailable.
Real-world studies for the assessment of medicinal products and medical devices[99]			
3A	Design consideration	Protocol/Statistical analysis plan	External control selection and analysis should be planned and specified a priori.
Methodology for the clinical development of medical devices[100]			
4A	Criteria for using an external control	Possibility of randomization	An externally controlled trial is acceptable in cases where a randomized comparison is not feasible or when direct comparative evidence is unavailable.
4B	Design consideration	Protocol/Statistical analysis plan	External control selection and analysis should be planned and specified a priori.

4C	Criteria for using an external control	Possibility of randomization	An externally controlled trial is acceptable in cases where a randomized comparison is not feasible or ethical.
4D	Design consideration, Analysis consideration	Analytic method, Comparability across arms, Protocol/Statistical analysis plan	External control selection and analysis should be planned and specified a priori. The analytic methods chosen should identify, assess and manage all sources of confounding and bias in both the internal and external arms.
4E	Design consideration	Comparability across arms	Consider the impact of time frame differences and comparability between the internal and external arms on result interpretation. Consider differences in patient characteristics at baseline.
4F	Design consideration, Analysis consideration	Additional analyses, Analytic method, Protocol/Statistical analysis plan	External control selection and analysis should be planned and specified a priori. The analytic methods chosen should identify, assess and manage all sources of confounding and bias in both the internal and external arms. Sensitivity analyses should be performed to assess impact of confounding factors/sources of bias.
4G	Criteria for using an external control	Possibility of randomization	An externally controlled trial is acceptable in cases where a randomized comparison is not feasible or ethical or when direct comparative evidence is unavailable.
NICE real-world evidence framework[103]			
5A	Design consideration	Fit for purpose data	Select external control data that will minimise differences in key study variables between the internal and external arms.
5C	Design consideration	Fit for purpose data	Where possible, consider the use of data from a previous clinical trial to create an external control arm.
5D	Design consideration	Protocol/Statistical analysis plan	Target trial emulation should be considered, if unfeasible, trade-offs should be discussed. Select external control data that will minimise differences in key study variables between the internal and external arms, including patient characteristics, data collection, healthcare systems and time periods.
5E	Design consideration	Comparability across arms	The same inclusion criteria should be applied to the internal and external arms.
5F	Design consideration	Fit for purpose data	Select external control data from a similar time frame and health care system, where possible.
5G	Design consideration	Pooling data sources	If multiple potential sources of external data are suitable, pooling should only be done in cases where there is limited heterogeneity in terms of coverage and data quality.
5I	Analysis consideration	Additional analyses	Bias analyses should be planned a priori and performed to assess impact of confounding factors/sources of bias.

5J	Design consideration	Comparability across arms	Differences in variable definitions and data collection across the internal and external arms should be reported.
E8 (R1): General considerations for clinical studies[102]			
6A	Criteria for using an external control, Design consideration	Comparability across arms, Disease	Disease progression should be well known and highly predictable to incorporate an external control. Source population for external control arm(s) should be comparable to internal treatment arm(s).
6B	Design consideration	Protocol/Statistical analysis plan	Prespecify external control data source and design elements in protocol (baseline eligibility, exposure definition, endpoints, approaches to minimizing missing data and other sources of bias, and analytic approach).
E10: Choice of control group and related issues in clinical trials[101]			
7A	Criteria for using an external control	Disease, Effect size	An external control is most appropriate when the anticipated treatment effect is large. Disease progression should be well known and highly predictable to incorporate an external control.
7B	Criteria for using an external control	Outcomes	Endpoints should be objective, and the impact of baseline and treatment variables well characterized.
7C	Design consideration	Comparability across arms, Fit for purpose data	Source population for external control arm(s) should be comparable to internal treatment arm(s). Where possible, select external control data with detailed individual patient data on baseline demographics/status, concomitant therapy and course on study.
7D	Design consideration	Comparability across arms, Outcomes	Timing and method of outcome assessment should be consistent across the internal and external arms.
7E	Design consideration	Protocol/Statistical analysis plan	External control selection and analysis should be planned and specified a priori.
7F	Design consideration	Protocol/Statistical analysis plan	Analytic method to identify and manage sources of confounding and bias should be identified a priori.
7G	Design consideration	Pooling data sources	Multiple controls can be studied when no single optimal external control exits, if pre-planned and analyzed and interpreted conservatively.
7H	Design consideration	Outcomes	Where possible, consider blinding for outcome assessment in the internal and external arms.
7I	Criteria for using an external control	Disease, Possibility of randomizations	Disease progression should be well known and highly predictable to incorporate an external control. An externally controlled trial is acceptable in cases where a randomized comparison is not feasible or ethical or when direct comparative evidence is unavailable.

7J	Criteria for using an external control, Design considerations	Comparability across arms, Disease, Effect size, Outcomes	Endpoints should be objective, and the impact of baseline and treatment variables well characterized. An external control is most persuasive when the anticipated treatment effect is large. Prognostic factors should be well understood and available in the external control data source. The external control should closely resemble the internal arms.
7K	Analysis consideration	Analytic method	Consider a conservative approach, including the potential for a more stringent statistical significance level, given the potential for bias in favour of the treatment.

Table SI 3. Characteristics of studies included in the prioritized subset of articles in the scoping review on externally controlled trials in pediatric populations with genetic rare diseases (n=18)

ID[a]	Characteristics linked to guidance statements	(Konstan et al., 2017)[53]	(Clemens et al., 2020)[57]	(Clemens et al., 2022)[61]	(Clemens et al., 2023)[64]	(Harper et al., 2024)[68]	(Khan et al., 2019)[56]	(McDonald et al., 2021)[60]	(Mah et al., 2022)[63]	(Zaidman et al., 2023)[67]	(Pignolo et al., 2023)[66]	(Fumagalli et al., 2022)[62]	(Hendriksz et al., 2016)[52]	(Schulz et al., 2018)[55]	(Schulz et al., 2024)[69]	(Krosschell et al., 2018)[54]	(Day et al., 2021)[59]	(Mercuri et al., 2023)[65]	(Muntoni et al., 2020)[58]
ROBINS-I V2 domain: Confounding																			
8	Individual patient information for external control	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
9	Same/highly similar inclusion criteria applied to the internal and external arms	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
10	Balance of baseline covariates compared	X	X	X	X	X	X	X	X	X	X	X	X			X		X	X
11	Applied analytic method to address differences in baseline covariates	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
12	Demonstrated the success of analytic method used to balance covariates	X							X										X
13	Conducted a bias analysis																		
ROBINS-I V2 domain: Outcome measurement																			
16	Same or highly similar outcome measurement instrument used for internal and external arms[b]	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
16	Same assessment time points across arms[b]	X	X	X	X					X							X	X	X
18	Blinding in assessment of at least one outcome										X								
ROBINS-I V2 domain: Selection of participants																			

19	Specified how the index date was selected for their external arm(s)[b]	X							X		X		X				X	X	
ROBINS-I V2 domain: Missing data																			
21	Clearly specified how missing data in the external arm(s) would be addressed in protocol	NA								NA	X	NA	NA			NA			
22	Reported missingness in external arm(s) baseline covariates		X				X				X					X			
22	Reporting missingness in external arm(s) outcome measurements		X	X	X				X		X	X	X	X	X				X
22	Reporting missingness in internal arm(s) baseline covariates		X				X				X					X			
22	Reporting missingness in internal arm(s) outcome measurements	X	X	X	X			X	X		X	X	X	X	X				X
Multiple ROBINS-I V2 domains																			
23	Pre-specified external control		X	X	X	X	X		X	X	X	X		X	X		X	X	X
25	Discussed target trial emulation																		
26	Discussed the estimand framework																		
27	External control data from clinical trial									X								X	
28	Multiple external controls					X		X	X			X	X						
29	Pooling external control data									X								X	X
30	Same or highly similar outcome measurement instrument used for internal and external arms[b]	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
33	Sensitivity analysis including external data	X					X		X					X	X				

[a]Corresponds to synthesized guidance IDs in Tables 3, 4, 5, and 6.
[b]For all efficacy and safety outcomes compared to an external control arm.

4. Integrated discussion

4.1 Introduction to the chapter

This thesis aimed to describe the design and conduct of externally controlled trials in pediatric populations with genetic rare diseases (RDs), and to critically evaluate current practices by comparison to existing regulatory and health technology assessment (HTA) guidance. We first conducted a scoping review to identify and describe the methodological and reporting characteristics of externally controlled trials in pediatric populations with genetic RDs (Chapter 2). We subsequently conducted a focused literature review to identify and synthesize current guidance specific to externally controlled trials published by leading regulatory bodies and HTA organizations and we compared the externally controlled pediatric RD trials to this guidance (Chapter 3). Overall, the findings of this thesis highlight areas for improvement for the future conduct of externally controlled trials in pediatric genetic RDs, including recommendations to mitigate potential biases.

4.2 Summary of findings

4.2.1 Chapter 2: The design, conduct and reporting of externally controlled trials in pediatric rare genetic disease

Following JBI methodology,[1] we conducted a scoping review and identified 67 articles that reported on an externally controlled trial conducted in pediatric populations with genetic RDs. The majority of these articles reported the results of a main trial (versus an extension study) and included a single internal intervention arm. The median sample size across internal arms was 22 participants. Sixteen articles (24%) described trials that made a comparison to aggregate external data, and 51 (76%) reported a comparison to individual-level external data. The design and use

of external data was heterogeneous across articles, as was the level of detail they included about the external comparison. We therefore decided to extract general information about all of the externally controlled trials that met our inclusion criteria (n=67 articles) and to extract more detailed information from a prioritized subset of 18 articles describing trials that incorporated individual-level external data and applied a specific analytic method in an attempt to address differences in baseline characteristics between the internal and external arms beyond age and baseline outcome score.

Among the prioritized subset, most articles pre-specified the incorporation of an external control arm, and all 18 included external data from at least one observational source, such as a cohort study, natural history study or longitudinal patient registry. A majority of the articles reported on trials of treatments for Duchenne muscular dystrophy (n=8) or types of spinal muscular atrophy (n=4). Several of the articles were linked manuscripts from the same trial program or used the same source of external control data. The baseline covariates considered as potential confounders and the methods for addressing confounding varied widely across articles. The majority of these articles reported comparing a primary trial outcome to external control data, yet reporting on outcome assessment timepoints in the external control data was often poor. Few articles reported on missingness in the external arm, the handling of missingness was rarely addressed, and sensitivity and bias analyses for external data were insufficiently implemented.

4.2.2 Chapter 3: Guidance for the conduct of externally controlled trials

We conducted a focused literature review to identify guidance available through the websites and repositories of 60 eligible regulatory and HTA organizations. We identified seven documents that provided guidance on the design, conduct and analysis of externally controlled trials. From

these, we extracted and synthesized all verbatim guidance specific to externally controlled trials, yielding a set of 33 final guidance statements that captured criteria for using an external control, design considerations, and/or analysis considerations for externally controlled trials. We subsequently mapped guidance to the domains of the Risk Of Bias In Non-randomized Studies – of Interventions, Version 2 (ROBINS-I V2) tool,[2] an existing risk-of-bias tool that may already be used by reviewers of externally controlled trials,[3] and by trialists or sponsors seeking to design such trials with mitigation of bias in mind, to explore their alignment or deviation. Our summarized guidance aligned with certain domains of the ROBINS-I V2 tool. Most guidance statements were either broad and corresponded to several ROBINS-I V2 domains or were found to be related to minimizing confounding by increasing the comparability of participants across the internal and external arms at both the design and analysis stages. Other recommendations included considerations for outcome assessment, missing data and sensitivity analyses. We used our synthesized guidance statements to assess current practices and reporting in the externally controlled pediatric genetic RD trials that were described in Chapter 2. From this assessment we identified important limitations across many of the reviewed trials, such as: lack of application of specific analytic methods to address baseline differences, unclear establishment of a comparable time at which follow up should begin in external control arms, limited reporting of missing data in external control arms, and poor utilization of additional analyses to demonstrate comparability and address concerns due to potential sources of bias.

4.3 Interpretation of findings in the context of related literature

Based on the findings presented in Chapters 2 and 3, the conclusions of many externally controlled trials in pediatric genetic RDs published over the last decade may be at risk of bias

due to differences between internal and external arms in baseline characteristics and outcome assessments, and due to poor handling of missing data. Furthermore, poor reporting practices and low uptake of sensitivity analyses call into question the reliability of the interpretation of results by trial report authors. These concerns do not appear to be specific to the field of pediatric genetic RDs. In their review of externally controlled trials in participants with immune-mediated inflammatory disease, Zayadi et al.[4] found that 79% of reviewed articles did not adjust for known covariates to improve comparability between arms and noted that among the studies reviewed, reporting of eligibility criteria, baseline characteristics, outcome definitions, covariates, and presence of missing data was often poor. Liu et al.[5] reviewed externally controlled trials across various disease areas and noted that only a third used a statistical method to adjust for important covariates across arms, approximately 37% did not specify inclusion criteria for the external arm, just 7% specified how missingness was to be addressed in the external control, and only 18% conducted a sensitivity analysis. These findings are further corroborated by a review of therapeutic submissions that utilized external controls, in which Sola-Morales et al. highlighted that regulatory and HTA organizations commonly raised concerns over the comparability of arms at baseline, choice of outcomes, and missing data.[6] In contrast, in their review of externally controlled trials in the field of hematological cancers, Hermans et al.[7] reported that 72% of studies applied matching and only 41% failed to report how missingness should be addressed in the external arm. The median sample size across internal arms was larger in the trials included in the review by Hermans et al.[7] (median = 127) in comparison to those included in the reviews conducted by Liu et al.[5] (median = 47), Zayadi et al.[4] (median = 73) and our review (median = 22). Smaller sample sizes may make some recommendations more challenging to implement, for example, regarding analytic approaches to

address baseline covariate imbalance. Thus, future guidance on the conduct of externally controlled trials should include specific and implementable recommendations for trials where small sample sizes are inevitable, such as in RDs.

One broad strategy recommended in our synthesized guidance statements was to apply causal inference frameworks and methods, such as target trial emulation and the estimand framework,[8,9] in the conception and design of externally controlled trials.[10,11] This may help to prevent or mitigate several of the above-mentioned biases by guiding design choices, for example regarding external control participant selection. Specifically, the clear pre-specification of an appropriate research question and associated key design elements for the internal and external arms, such as the inclusion and exclusion criteria, treatment strategies, assignments, outcomes of interest, follow-up, and chosen statistical approaches are recommended. Among our prioritized subset of articles reporting on externally controlled trials, identified in the scoping review in Chapter 2, there was no discussion of target trial emulation nor the estimand framework and how these influenced the design of the trials or interpretation of treatment effects. This introduced a key methodological gap in several studies, however, we note that several of the trials were published prior to the availability of the estimand framework and prior to the recent widespread uptake of target trial emulation approaches. We recommend the use of causal inference methods in the design of externally controlled trials moving forward, echoing recommendations made by Liu and colleagues.[5]

4.3.1 Risk of bias due to confounding in externally controlled pediatric genetic RD trials and opportunities for better practice

In randomized controlled trials (RCTs), random allocation promotes comparability between groups at baseline.[12] Given that it is not possible to randomly allocate a participant to the internal or external arm, demonstrating comparability between arms is essential in an externally controlled trial.[13] Comparability should be considered at both the design and analysis stages. As highlighted in Chapter 3, where possible, trialists should select external data with a source population similar to the investigational arm and apply the same inclusion and exclusion criteria across arms to limit differences in key study variables at baseline. Prior to the conduct of the externally controlled trial, comprehensive comparability assessments of potential confounders and prognostic factors known at baseline should be conducted and clearly reported. However, this is only possible when using data sources with sufficient patient-level information.

Specific analytic methods to increase the comparability of internal and external arms at baseline should be pre-specified and applied prior to the interpretation of study data. In Chapter 3 we identified a lack of detailed guidance from regulatory and HTA organizations regarding which specific analytic methods may best increase the comparability of study arms in externally controlled trials. In their recent review, Hogervorst et al.[14] identified a range of analytic methods from the scientific literature specific to comparability of arms in externally controlled trials, including methodologically-focused case studies and articles focused on the development of new methods. They compared these methods from the scientific literature to the approaches discussed in guidelines or articles from the European Medicines Agency and four European HTA organizations and similar to our findings, concluded that there is a need for additional guidance

specific to particular research questions and settings.[14] This should include considerations for small sample sizes, to support the conduct of externally controlled trials in RD populations.

4.3.2 Time-related biases in externally controlled pediatric genetic RD trials and opportunities for better practice

In RCTs, biases arising from time-related differences between arms are commonly mitigated because the selection of the time at which follow-up begins (time zero), and the calendar time in which they are followed is typically the same across all arms.[15] In RCTs, time zero is commonly the point at which individuals are randomized and is the same across all participants, subsequently avoiding several time-related biases such as immortal time bias and survivorship bias. Immortal time is a time span of follow-up time during which a participant, by definition, could not experience the event of interest. Immortal time bias can occur when time zero and allocation to an exposure category or receipt of an exposure or intervention do not align, resulting in biased estimates.[16] Survivorship bias, a form of selection bias, is introduced when analyses are focused on individuals who have already survived a certain amount of time, while excluding those who did not.[17] Both immortal time bias and survivorship bias can result in distorted estimates of treatment effects. Selecting an appropriate and comparable time zero can be challenging for the external arm in externally controlled trials. For example, participants followed in a natural history study may meet the eligibility criteria for the internal trial arm at multiple time points. In our prioritized subset of externally controlled pediatric RD trials from the scoping review, we noted that time zero was often unspecified in the external arm, making it challenging to evaluate the risk of time-related bias. To demonstrate that the risk of bias related to time zero was carefully considered and mitigated, trialists should clearly pre-specify and

justify the appropriateness and comparability of the selected time zero across all arms. In cases where more than one time zero may be appropriate, trialists may consider conducting sensitivity analyses to investigate whether the results are robust to the selection of different appropriate time points.[18,19] Externally controlled trials that incorporate a non-concurrent control may also at risk of time-related bias due to differences in calendar time, as the comparator treatment or standard of care in the historical cohort may have improved over time.[15] The impact of calendar-time differences should this also be considered by trialists when interpreting their results.[20] Specific guidance is needed to support trialists in the mitigation of this time-related bias when an appropriate concurrent control is not available.

In the synthesized guidance presented in Chapter 3, organizations recommended that outcome measurement methods and timing should be consistent across the internal and external arms.[11,20] However, when choosing to compare a repeated measures or time-to-event outcome, this may be difficult to achieve, and indeed some of the articles in our scoping review compared outcomes measured at non-equivalent time points across arms. Regulators have commonly expressed concerns over the use of time-to-event endpoints in externally controlled trials due to limited overlap in timing of outcome assessment across arms, making the results difficult to interpret.[6] Comparing outcomes with different assessment time points may introduce measurement bias and limit comparability of the internal and external arms. Having outcome assessment similarly distributed through time would be considered best practice but may not be possible when using real-world data (RWD), where outcome assessments may be dictated by clinic visit timepoints. Conducting sensitivity and quantitative bias analyses to assess the extent to which the observed effect of treatment may be biased may further support the interpretation of the treatment effects of externally controlled trials. Where possible, trialists could consider the

use of external data from a clinical trial (e.g., from a placebo arm),[10,11] whereby outcome assessments were protocolized, as this may help decrease the risk of measurement bias inherent in RWD where assessment time points correspond to clinical needs rather than protocolized research assessments. Nonetheless, further guidance is needed on the selection and evaluation of outcomes when incorporating RWD.

4.3.3 Risks of bias due to missingness in externally controlled pediatric genetic rare disease trials and opportunities for better practice

Missing data are of high concern in externally controlled trials, due to the lack of protocolization of data collection in many observational data sources that could be used as an external control arm. The pre-specification of methods to address missingness in the external control arm and reporting of missingness in externally controlled trials was found to be low in our prioritized subset of articles from the scoping review in Chapter 2. This aligns with the findings of recent reviews of externally controlled trials in other disease areas.[4,5] Trialists should ensure missingness in the external control arms is assessed and reported prior to the conduct of an externally controlled trial. Further to this, methods to address missingness should be pre-specified and their assumptions tested and shown to be met through additional analyses, for example, by conducting a tipping point analysis.[21]

4.3.4 Data quality in externally controlled pediatric genetic RD trials

Given that observational data are not always collected in a protocolized manner (as distinct from clinical trial data), the quality of external control data, including their accuracy, completeness, and consistency, is central to the mitigation of potential biases and thus the credibility of external

comparisons.[10,22-27] Observational data from natural history studies and longitudinal patient registries were used as a source of external control data in all of the articles included in our prioritized subset of pediatric RD trials from the scoping review in Chapter 2. Three articles additionally included routinely collected data and data from previously conducted clinical trials. We noted that one natural history cohort (the Cooperative International Neuromuscular Research Group Duchenne Natural History Study) was utilized as a source of external control data across all articles in that prioritized subset that included patients with Duchenne muscular dystrophy, regardless of treatment.[28] This natural history study was created for the purpose of informing future clinical trials and follows a protocol with an assessment schedule for data collection, which includes assessments at baseline and pre-specified follow-up visits. Further development of observational research studies and registries that thoughtfully consider the guidance statements presented in the Chapter 3 and include protocolization as in this example, may positively impact the conduct and acceptability of future externally controlled trials. The protocolization of other sources of RWD, such as insurance claims data and hospital records may not be possible due to the purpose of data collection. Nonetheless, selecting a data source that has demonstrated to be of high quality for the variables of interest, may help mitigate several sources of bias.

Taken together, the findings of this thesis highlight a gap between current practices in pediatric genetic RD externally controlled trials and guidance offered by regulatory and HTA organizations. The results also underscore a need for more specific guidance for addressing risks of bias due to confounding, measurement of outcomes, selection of time zero and missing data. Trialists should ensure the pre-specification and transparent reporting of all design and analysis

elements, including testing all assumptions and conducting sensitivity analyses to justify decisions.

4.4 Strengths and limitations

This research has several notable strengths. We systematically conducted a scoping review using established methodology,[1] including developing and publishing a protocol a priori,[29] screening by two team members working independently, and verification of data extracted, to ensure research rigor. We also synthesized currently available regulatory and HTA guidance on the design, conduct and analysis of externally controlled trials, which addressed a gap by summarizing current recommendations about this trial design across relevant policy organizations. This summary of the current guidance available to trialists allowed us to identify areas for further consideration and better practice in future externally controlled pediatric genetic RD trials as well as identify gaps in currently available guidance. Researchers with expertise in pediatric rare disease, externally controlled trials, and regulatory frameworks for drug approval were involved in the conception, design, and interpretation of findings, strengthening the overall quality and reliability of this research project's conclusions.

Nonetheless, several limitations should be considered when interpreting the results of this research. As noted in Chapter 2, all articles reviewed for our scoping review were limited to those published and are therefore subject to potential publication bias. Nonetheless, our scoping review aimed to characterize the design and methodology of these trials, rather than the results. Similarly, as discussed in Chapter 3, we only considered publicly available regulatory and HTA documents in our focused review, thus any additional guidance for externally controlled trials available within organizations but not publicly posted was not included in the synthesized

guidance presented. In our focused review, we aimed to consider documents from a comprehensive set of regulatory and HTA organizations. In an effort to do this systematically, we decided to limit guidance documents to those published by founding and standing regulatory members of the International Council for Harmonization of Technical Requirements for Pharmaceuticals for Human Use (ICH), as well as all HTA organizations that are members of International Network of Agencies for Health Technology Assessment. However, other organizations that were not considered may have published guidance specific to externally controlled trials. In addition, other eligible externally controlled trials or guidance documents may have been published since the search dates of our literature reviews. For example, the United States Food and Drug Administration (FDA) published a draft guidance document in the months following our website and guidance repository search that includes guidance on externally controlled trials in small populations.[30] However, much of the guidance included in this document was also included in the FDA's *Considerations for the design and conduct of externally controlled trials for drug and biological products guidance for industry: Draft guidance,*[11] which was included in our guidance synthesis. Finally, our scoping review was limited by the reporting of trial authors. Despite considering all publicly available documents linked to the study reports during extraction, such as trial registrations, protocols, statistical analysis plans, and supplemental material, certain data fields were missing for some articles. For example, time zero in the external control arm was poorly reported across several articles in our prioritized subset. This information could have been included in the trialist's submission to regulatory and HTA organizations for approval and reimbursement, respectively, but in this review, it was considered unreported or unspecified based on published reports.

4.5 Implications for future research, practice, and policy

To build on the findings of this research, future studies should explore the data quality of RWD sources utilized in externally controlled pediatric genetic RD trials. This information may be valuable not only for trialists in pediatric genetic RD in the selection of their external control data sources but also for RWD producers in robustly designing longitudinal data collection to support future externally controlled trials. Furthermore, trialists, regulatory agencies and HTA organizations may benefit from the development of a risk-of-bias tool that is specific to externally controlled trials and includes considerations for small sample sizes.

More specific guidance is needed to ensure trialists are provided with implementable recommendations for best practice prior to the conduct of externally controlled trials. Collaboration across several regulatory or HTA organizations to create harmonized guidance for RDs may increase relevance and uptake (e.g., ICH). This guidance should emphasize recommendations to avoid biases related to confounding, on-study time, outcome assessment and missing data, while taking into consideration their feasibility for RD populations. In the interim, trialists conducting externally controlled trials should carefully consider the synthesized guidance presented in our third chapter when designing and analyzing externally controlled trials to mitigate biases where possible. The collaborative development of reporting guidelines and a risk-of-bias tool specific to externally controlled trials could also be beneficial in ensuring uniformity and transparency in reporting practices for externally controlled trials across disease areas and countries.

4.6 Conclusions

Given that RCTs are often unfeasible or unethical to conduct in pediatric populations with genetic RDs, externally controlled trials are a promising alternative study design to assess the efficacy and comparative effectiveness of potentially disease-modifying therapies. This thesis aimed to support the conduct of future externally controlled trials in pediatric genetic RD by identifying and describing the design, methodological, and reporting characteristics of those published over the last decade and by synthesizing the current regulatory and HTA guidance on the conduct of externally controlled trials. Our review highlighted substantial heterogeneity in the use of external controls and their reporting across studies. The comparison of these trials to current guidance pointed to the potential for an elevated risk of bias among many published externally controlled trials in pediatric populations with genetic RDs. We discussed the design and analysis considerations for mitigating these common biases and highlighted areas in which more guidance is needed to support trialists.

The findings of this thesis contribute to the ongoing discussion about the reliability of externally controlled trials, specifically for therapies indicated for pediatric populations with genetic RDs. The findings and works presented in this thesis will be an important tool for trialists in the conduct of more robust externally controlled trials in pediatric genetic RD, thereby supporting the generation of evidence for potentially disease-modifying therapies in a population with a high unmet need and where randomization is often considered infeasible.

4.7 References for Chapter 4

1. Peters MDJ, Marnie C, Tricco AC, et al. Updated methodological guidance for the conduct of scoping reviews. *JBI Evid Implement*. 2021;19(1):3-10. doi:10.1097/xeb.0000000000000277

2. Sterne J, Brandt Mathur M, Elbers R, et al. The Risk Of Bias In Non-randomized Studies – of Interventions, Version 2 (ROBINS-I V2) - Draft. Accessed April 1, 2026. https://drive.google.com/file/d/1tXCLVa7aFVaiV3OnitZDgAnX3kXWXor0/view

3. Canada's Drug Agency. *Methods guide for health technology assessment*. 2025. Accessed June 16, 2025. https://www.cda-amc.ca/sites/default/files/MG%20Methods/MG0030-Quantitative-Methods-Manual_Mar%202025.pdf

4. Zayadi A, Edge R, Parker CE, et al. Use of external control arms in immune-mediated inflammatory diseases: A systematic review. *BMJ Open*. Dec 9 2023;13(12):e076677. doi:10.1136/bmjopen-2023-076677

5. Liu J, Yao M, Wang M, et al. Design, conduct, and analysis of externally controlled trials. *JAMA Netw Open*. 2025;8(9):e2530277-e2530277. doi:10.1001/jamanetworkopen.2025.30277

6. Sola-Morales O, Curtis LH, Heidt J, et al. Effectively leveraging RWD for external controls: A systematic literature review of regulatory and HTA decisions. *Clin Pharmacol Ther*. 2023;114:325-355. doi:10.1002/cpt.2914

7. Hermans SJF, van der Maas NG, van Norden Y, et al. Externally controlled studies using real-world data in patients with hematological cancers: A systematic review. *JAMA Oncol*. 2024;10(10):1426-1436. doi:10.1001/jamaoncol.2024.3466

8. Hernán MA, Dahabreh IJ, Dickerman BA, Swanson SA. The target trial framework for causal inference from observational data: Why and when is it helpful? *Ann Intern Med*. 2025;178(3):402-407. doi:10.7326/annals-24-01871

9. International Council for Harmonization of Technical Requirements for Pharmaceuticals for Human Use. *E9(R1): Addendum on estimands and sensitivity analysis in clinical trials to the guideline on statistical principles for clinical trials*. 2019. Accessed May 1, 2026. https://database.ich.org/sites/default/files/E9-R1_Step4_Guideline_2019_1203.pdf

10. National Insitute for Health and Care Excellence. *NICE real-world evidence framework*. 2022. Accessed June 16, 2025. https://www.nice.org.uk/corporate/ecd9/resources/nice-realworld-evidence-framework-pdf-1124020816837

11. United States Food and Drug Administration. *Considerations for the design and conduct of externally controlled trials for drug and biological products guidance for industry: Draft guidance*. 2023. Accessed October 24, 2024. https://www.fda.gov/vaccines-blood-biologics/guidance-compliance-regulatory-information-biologics/biologics-guidances

12. Lim CY, In J. Randomization in clinical studies. *Korean J Anesthesiol*. 2019;72(3):221-232. doi:10.4097/kja.19049

13. Seeger JD, Davis KJ, Iannacone MR, et al. Methods for external control groups for single arm trials or long-term uncontrolled extensions to randomized clinical trials. *Pharmacoepidemiol Drug Saf*. 2020;29(11):1382-1392. doi:10.1002/pds.5141

14. Hogervorst MA, Soman KV, Gardarsdottir H, Goettsch WG, Bloem LT. Analytical methods for comparing uncontrolled trials with external controls from real-world data: A systematic literature review and comparison with European regulatory and health technology assessment practice. *Value Health*. 2025;28(1):161-174. doi:10.1016/j.jval.2024.08.002

15. Suissa S. Single-arm trials with historical controls: Study designs to avoid time-related biases. *Epidemiology*. 2021;32(1):94-100. doi:10.1097/ede.0000000000001267

16. Hernán MA, Sterne JAC, Higgins JPT, Shrier I, Hernández-Díaz S. A structural description of biases that generate immortal time. *Epidemiology*. 2025;36(1):107-114. doi:10.1097/ede.0000000000001808

17. Elston DM. Survivorship bias. *J Am Acad Dermatol*. 2021;(21):01986-01991. doi:10.1016/j.jaad.2021.06.845

18. Ackerman B, Gan RW, Meyer CS, et al. Measurement error and bias in real-world oncology endpoints when constructing external control arms. *Front Drug Saf Regul*. 2024;4:1423493. doi:10.3389/fdsfr.2024.1423493

19. Ackerman B, Gan RW, Zhang Y, et al. Quantitative bias analyses to address measurement error in time-to-event endpoints. *Am J Epidemiol*. 2026;doi:10.1093/aje/kwag027

20. International Council for Harmonization of Technical Requirements for Pharmaceuticals for Human Use. *E10: Choice of control group and related issues in clinical trials*. 2000. Accessed June 16, 2025. https://database.ich.org/sites/default/files/E10_Guideline.pdf

21. Liu Y, Zhou K, Sims KD. Tipping point analysis: Assessing the potential impact of missing data. *JAMA*. 2025;334(3):265-266. doi:10.1001/jama.2025.7420

22. United States Food and Drug Administration. *Real-world data: Assessing registries to support regulatory decision-making for drug and biological products guidance for industry*. 2023. Accessed May 25, 2026. https://www.fda.gov/media/154449/download

23. European Medicines Agency. *Guideline on registry-based studies*. 2021. Accessed May 25, 2026. https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-registry-based-studies_en.pdf

24. Gatto NM, Campbell UB, Rubinstein E, et al. The structured process to identify fit-for-purpose data: A data feasibility assessment framework. *Clin Pharmacol Ther*. 2022;111(1):122-134. doi:10.1002/cpt.2466

25. Haute Autorité de Santé. *Real-world studies for the assessment of medicinal products and medical devices*. 2021. Accessed June 16, 2025. https://www.has-sante.fr/upload/docs/application/pdf/2021-10/real-world_studies_for_the_assessment_of_medicinal_products_and_medical_devices.pdf

26. United States Food and Drug Administration. *Use of real-world evidence to support regulatory decision-making for medical devices: Guidance for industry and Food and Drug Administration staff*. 2025. Accessed May 25, 2026. https://www.fda.gov/downloads/medicaldevices/deviceregulationandguidance/guidancedocuments/ucm513027.pdf

27. Daniel G, McClellan M, Silcox C, Romine M, Bryan J, Frank K. *White Paper: Characterizing RWD quality and relevancy for regulatory purposes*. 2018. Accessed May 25, 2026. https://healthpolicy.duke.edu/sites/default/files/2020-03/characterizing_rwd.pdf

28. McDonald CM, Henricson EK, Abresch RT, et al. The cooperative international neuromuscular research group Duchenne natural history study - A longitudinal investigation in the era of glucocorticoid therapy: Design of protocol and the methods used. *Muscle Nerve*. 2013;48(1):32-54. doi:10.1002/mus.23807

29. Pratt Tremblay G, Iverson E, Bliu A, Gravel C, Potter B. Protocol 3: Externally controlled clinical trials for pediatric rare disease: Current practices and guidance. *OSF*. 2025. doi:10.17605/OSF.IO/4MBRN

30. United States Food and Drug Administration. *Innovative designs for clinical trials of cellular and gene therapy products in small populations: Draft guidance for industry*. 2025. Accessed November 15, 2025. https://www.fda.gov/media/188892/download