RTI (h)(s) Health Solutions

Use of External Control Arms in Marketing Applications Submitted to European Medicines Agency: A Targeted Review

MSR60

Katherine Houghton, Anubhav Nangia, Emma Hawe

RTI Health Solutions, Manchester, United Kingdom

BACKGROUND

- For some conditions, research with fewer patients involved is justified due to the rarity of the condition. Additionally, lack of exposure to a control may be justified due to ethical concerns associated with assigning very sick patients to placebo or standard of care
- Single-arm trials are increasingly reported as the sole basis for efficacy by the US Food and Drug Administration (breakthrough therapy designation) and the European Medicines Agency (EMA) (accelerated access)[1-4]
- External control arms (ECAs) are increasingly being utilised to estimate comparative treatment effect in the absence of an appropriate within-trial comparative arm[5-7]

OBJECTIVE

- To examine the frequency with which recent applications submitted to EMA include data from ECAs, associated therapeutic areas, methods used to derive ECAs, and feedback received regarding the ECAs

METHODS

- The EMA website was searched for European Public Assessment Reports (EPARs) published from 1 January 2022 through 23 May 2023
- The following EPARs were excluded: medical devices, vaccines, diagnostic technologies, hybrid medicines, and biosimilars. The clinical efficacy sections in EPARs of other products were reviewed, and these EPARs were searched using the terms "external control," "observational," and "real world"
- Data were extracted from EPARs in which ECAs were mentioned

RESULTS

- 54 EPARs met the study inclusion criteria; of these, 11 (20%) included an ECA
- The identified EPARs primarily assessed products for rare/ultra-rare diseases (9/11; 82%). The products were assessed mainly in single-arm trials (SATs) (9/11; 82%), with 2 products investigated in randomised controlled trials (RCTs) (2/11; 18%)
- ECAs were derived from retrospective medical record reviews, prospective observational studies, historical trials, clinical databases, and literature reviews (Table 1)
- Matching techniques or propensity score weighting analyses to increase comparability between patients in the ECAs and experimental trials were applied in 7 EPARs, and naïve comparisons without matching or weighting were used in 4 EPARs (Table 1)
- The use of ECAs was considered acceptable by the EMA only under the following circumstances:
  - When no other comparative data source exists due to the rarity and severity of the condition
  - When improvements associated with the active treatment are consistent with those observed in a similar population for which a comparative RCT was performed (e.g., SAT with ECA in children, comparative RCT in adults)
- Other aspects considered favourable by the EMA included prespecification of the use of an ECA and appropriate matching or adjusting techniques with key characteristics considered
- The EMA's primary concerns with the use of ECAs were selection bias and unmeasured confounding

DISCUSSION

- Our review found that ECAs are used to generate comparative treatment effects for EPARs primarily for rare/ultra-rare diseases where sample sizes are very small and data sources are limited
- Certain practices increase the acceptability of the use of ECAs, including careful consideration of selection bias with the ECA, use of statistical analytic approaches to reduce the potential impact of selection bias, and prespecification of analysis including ECA
- A key concern related to ECAs derived from real-world data is the impact of unmeasured confounding, which cannot be taken into consideration analytically, and this impact leads to uncertainty about the actual treatment effect

CONCLUSIONS

The EMA considers evidence submitted from ECAs acceptable when there is a lack of any other data source due to disease rarity, when bias in the ECA is considered not impactful to overall conclusions, and when population characteristics between the clinical trial and ECA are comparable

Table 1. Methods and Feedback Associated With ECAs in EMA EPARs

Product, Indication, Trial	ECA Source, Comparator Treatment(s)	Methods Used to Derive and Analyse ECA	Key Feedback
ECA acceptable			
Xenpozyme (olipudase alfa) • Acid sphingomyelinase deficiency • DFI13803 Peds; SAT of 20 patients	Prospective natural history study (MSC12840) of 14 patients	• Naïve comparison • ANCOVA models tested for a difference between the 2 groups adjusting for confounders	✓ Imbalance between the 2 studies was considered ✓ Showed relevant clinical improvements consistent with the results from the RCT conducted among adults
Nulibry (fosdenopterin) • Molybdenum cofactor deficiency type A • 3 SATs involving 15 patients	**37 untreated patients derived from:** • Retrospective medical record review • Prospective natural history cohort	• 1:many matching based on genotype (19 patients identified) • Average treatment effects on the treated weighted analysis	✓ Comparison is not problematic because it is the only available data source ✓ Selection bias was considered unlikely ✕ It is unclear whether matching based only on genotype captures all possible confounding between treatment and outcome
Amvuttra (vutrisiran) • Polyneuropathy caused by hereditary transthyretin-mediated amyloidosis • HELIOS-A; randomised open-label trial of 164 patients with an active comparator	External placebo group of 77 patients from APOLLO, an RCT with the same active comparator as HELIOS-A	• Propensity score used as a covariate in analytic models	✓ The use of ECA including a reference comparator helped mitigate concerns about the open-label nature of HELIOS-A ✓ Analyses prespecified ✓ Baseline disease characteristics were adjusted for
Zokinvy (lonafarnib) • Hutchinson-Gilford progeria syndrome and processing-deficient progeroid laminopathies • 2 SATs (Prolon1 and Prolon2) of 62 patients	Natural history control based on: • Progeria Research Foundation International Registry (211 patients) • Scientific publications (47 patients)	• Propensity score–based 1:1 matching, defining time 0 for controls as age of starting treatment for the matching partner	✓ Comparison is not problematic because it is the only available data source ✓ The matching algorithm is acceptable ✕ Patients were excluded from the natural history cohort due to health/mobility, which introduces selection bias ✕ The lack of understanding about the impact of baseline and treatment variables on survival makes a comparison challenging ✕ Comparison will unlikely overcome deficits of the SAT results not being able to determine whether a factor is prognostic or predictive of outcome
ECA supportive			
Carvykti (ciltacabtagene autoleucel) • Relapsed/refractory multiple myeloma • CARTITUDE-1; SAT of 113 patients	Retrospective medical record review of 275 patients	• Propensity score–based 1:1 matching • Stabilised inverse probability treatment weighting • Average treatment effects on the treated weighted analysis	✓ Showed relevant clinical improvements ✕ Due to the missing randomised control group, uncertainty about the actual treatment effect exists
Upstaza (eladocagene exuparvovec) • Severe aromatic L-amino acid decarboxylase deficiency • 3 SATs (AADC-CU/1601, AADC-010, AADC-011) of 26 patients	Natural history control of 82 patients based on SLR	• Data from severe phenotype patients who were fully dependent • Naïve comparison	✓ Reflects worldwide experience and consensus about the clinical presentation and diagnosis ✕ Unclear whether key characteristics match between the active treatment and patients in the ECA ✕ Not considered robust enough for statistical comparison, particularly due to sample size
Ebvallo (tabelecleucel) • Epstein-Barr virus–positive posttransplant lymphoproliferative disease • 1 SAT (ALLELE) of 43 patients	Retrospective medical record review of 302 patients	• Propensity score matching	✕ Considered supportive, considering well-known limitations of retrospective studies compared with RCT
Livmarli (maralixibat chloride) • Cholestatic pruritis caused by Alagille syndrome • ICONIC; placebo crossover trial of 31 patients plus an SAT of 8 patients	Natural history cohort of 84 patients from a global clinical research database	• Controls were selected based on prespecified criteria similar to those from the sponsor's studies • Naïve indirect comparison	✕ Selection bias: key data not a major inclusion criteria for the ECA ✕ Unknown factors may have contributed to the results ✕ Differential follow-up times: concern of bias
Nexviadyme (avalglucosidase alfa) • Pompe disease • Study EFC14028 (COMET); RCT of 100 patients with an active comparator	Treatment-naïve patients pooled with patients randomised to receive comparator treatment in COMET from: • A historical RCT of 60 patients • An open-label extension safety study of 19 patients	• Naïve indirect comparison • Baseline characteristics are said to be adjusted for, but no information is provided on the statistical methods used to perform these analyses	✕ Analyses were not preplanned ✕ The presence of bias with the use of ECA data is likely and difficult to quantify
Breyanzi (lisocabtagene maraleucel) • Relapsed/refractory diffuse large B-cell lymphoma, high-grade B-cell lymphoma, primary mediastinal large B-cell lymphoma, and follicular lymphoma grade 3B • 2 SATs in > 300 patients	• Systematic literature review • Retrospective medical record review of 407 patients	• Matching-adjusted indirect comparison using SLR data • Stabilised inverse probability treatment weighting for medical record data	✓ Provides data to support external validity and generalisability of SAT results ✕ Significant clinical and biological heterogeneity of aggressive large B-cell lymphomas limits the use of the data for indirect comparison ✕ Analyses and covariates for propensity score model not prespecified; different starting timepoints from different studies not taken into account
ECA non-supportive			
Tabrecta (capmatinib) • Advanced non-small cell lung cancer with cells that have a particular genetic mutation • GEOMETRY, an SAT of 100 patients	Natural history control based on: • Retrospective medical record review of 211 patients • Flatiron Health clinic-genomic database of 61 patients	• Propensity score weighting using the Flatiron data	✕ Important limitations that preclude use as an ECA: small sample size, efficacy assessments in the real world not systematically collected at prespecified times; RECIST 1.1 criteria not commonly used, selection bias (posttreatment assignment of cohorts), treatment and biomarker testing landscape changed, unobserved confounding

ANCOVA = analysis of covariance; RECIST = response evaluation criteria in solid tumors; SLR = systematic literature review.

REFERENCES

1. Kim C, Prasad V. JAMA Int Med. 2015 Dec 1;175(12):1992-4.
2. Davis C, et al. BMJ. 2017 Oct 4;359.
3. Hatswell AJ, et al. BMJ Open. 2016 Jun 1;6(6):e011666.
4. Salcher-Konrad MA, et al. Milbank Quart. 2020 Dec;98(4):1219-56.
5. Flynn R, et al. Clin Pharmacol Ther. 2022 Jan;111(1):90-7.
6. Wang X, et al. J Cancer Policy. 2023 Jan 14:100403.
7. Naumann-Winter F, et al. Frontiers Pharmacol. 2022 Aug 11;13:920336.

CONTACT INFORMATION

Katherine Houghton, MSc
Director, Health Economics

RTI Health Solutions
Didsbury
United Kingdom

Telephone: +44(0)161.446.1733
Email: khoughton@rti.org

The power of **knowledge.**
The value of **understanding.**

Presented at: ISPOR Europe 2023; 12-15 November 2023; Copenhagen, Denmark