Supplementing single-arm trial data in NICE submissions — are synthetic approaches leading the charge?

Caroline Solon, MSc[1]; Michaela Lysakova-Ivanova, MSc[2]
[1]Avalere Health, San Francisco, CA, USA; [2]Avalere Health, London, UK

Avalere Health™
EVERY PATIENT POSSIBLE.

HTA305

Introduction

- Randomized controlled trials (RCTs) are accepted as the gold standard for demonstrating safety and efficacy, but single-arm trials (SATs) are increasingly used to support health technology assessment (HTA) submissions, particularly in oncology and rare diseases.
- Synthetic control arms (SCAs) are an innovative type of external control, defined as cohorts of patients from external data sets (separate clinical trials or real-world data [RWD]) adjusted using statistical methods.[1] Within the literature, terminology use varies, with the terms "synthetic control" and "synthetic data" sometimes being used synonymously for both adjusted individual patient-level data (IPD) and simulated data.[2] Other analyses may consider "external" and "synthetic" controls to be distinct or the same.[3]
- SCAs have the potential to innovate clinical trial design by reducing the need for enrolling control cohorts. This could support trials in rare diseases, which may struggle with trial recruitment, mitigate ethical considerations where there is no clinical equipoise, and reduce the time for innovative treatments to reach patients.
- Using SCAs to support clinical evaluation is a paradigm shift in evidence-based medicine.[1] While a 13-fold increase in SAT-based HTA submissions was reported up to 2019,[3] the real-world application of synthetic controls in payer submissions remains unclear.

Objectives

- To explore the methods used to supplement SAT data in HTA submissions in England.
- To understand the impact of these methods on HTA outcomes.

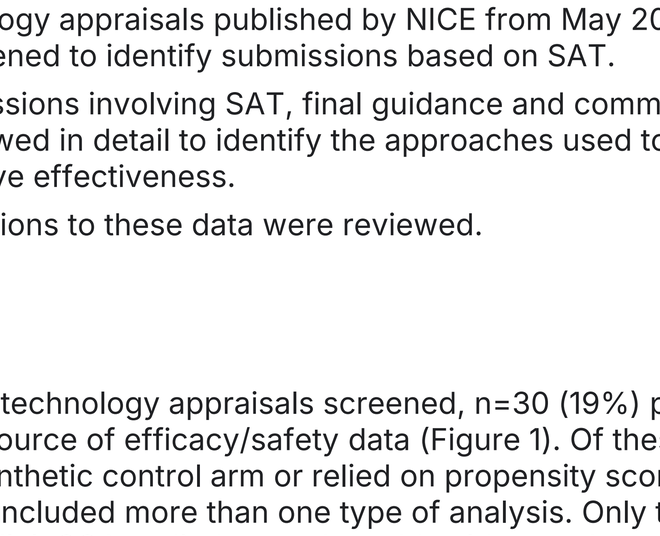
Methods

- All technology appraisals published by NICE from May 2023 to May 2025 were screened to identify submissions based on SAT.
- For submissions involving SAT, final guidance and committee papers were then reviewed in detail to identify the approaches used to provide comparative effectiveness.
- NICE reactions to these data were reviewed.

Results

Of the n=156 technology appraisals screened, n=30 (19%) presented SATs as the primary source of efficacy/safety data (Figure 1). Of these, n=7 explicitly involved a synthetic control arm or relied on propensity score analysis; most submissions included more than one type of analysis. Only two appraisals including explicit SCAs relied on real-world evidence alongside the pivotal SAT (Figure 2). Clinical inputs for comparator arms were usually derived from clinical trials using statistical methods, such as propensity score matching. Appraisals that did not include synthetic control arms mostly relied on statistical adjustments to allow matching-adjusted indirect comparisons (MAICs), naïve indirect treatment comparisons (ITCs), or both (Figure 3). No appraisals explicitly mentioned the use of machine learning (ML) algorithms or artificial intelligence (AI).

Figure 1: **Main source of efficacy data in submissions**



"Other" includes n=19 cost-comparison appraisals; n=2 combination of "living" systematic reviews and network meta-analyses (NMAs) for COVID-19 treatments; n=2 non-randomized ± randomized open-label trials; n=1 had been replaced by new guidance.

Trials with multiple sources including an RCT were classified as RCT, unless key subpopulations relied on SAT data.

Figure 2: **Sources of additional data for comparative efficacy in submissions with SATs as the primary source of efficacy data**

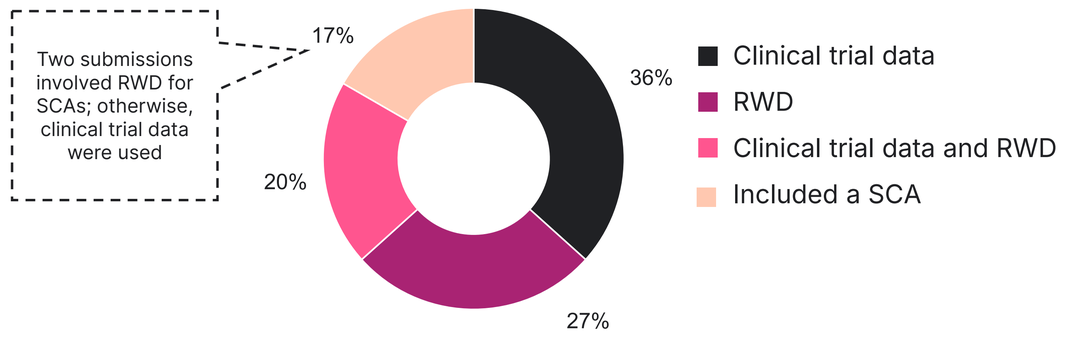
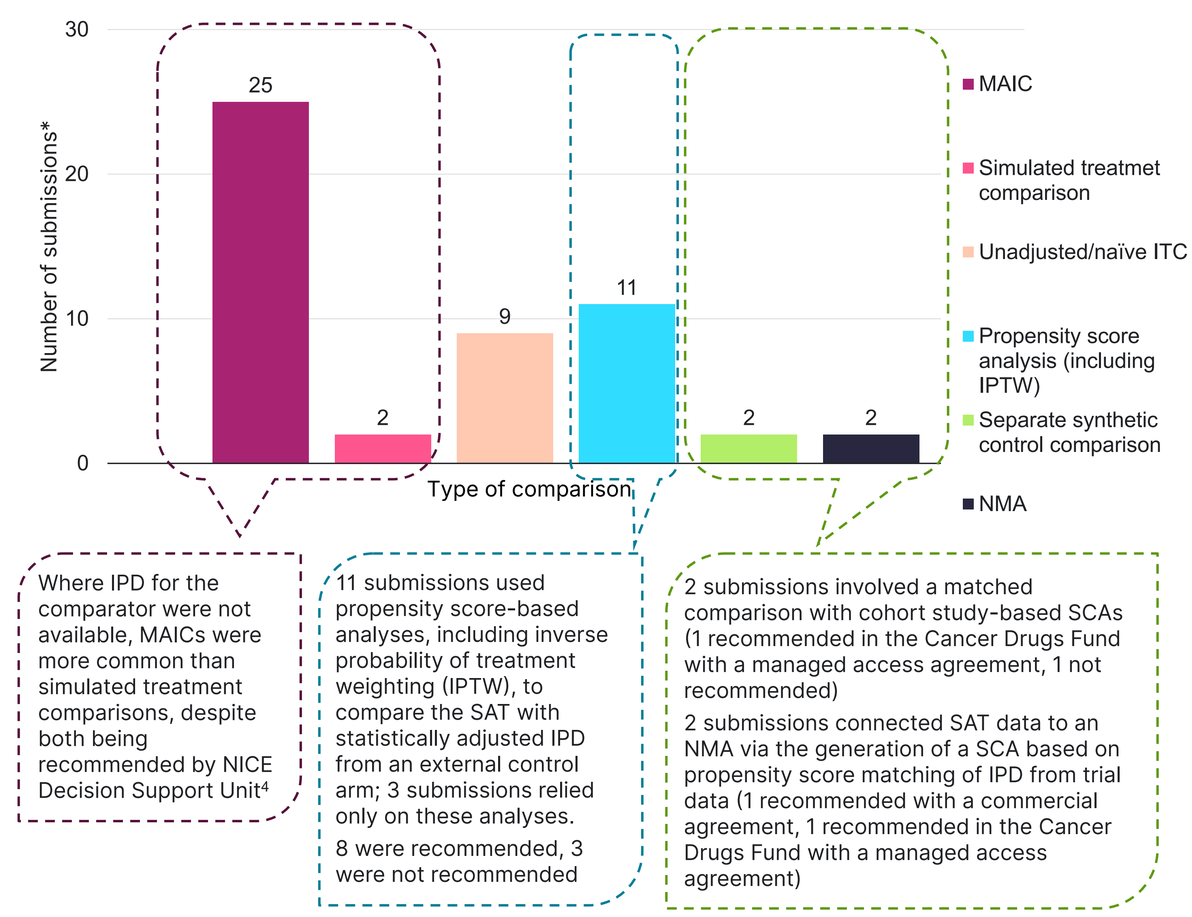


Figure 3: **Types of comparison in submissions with SATs as the primary source of efficacy data**



*As submissions could have more than one type of comparison, the number of submissions in the graph exceeds the total number of appraisals with SATs as the primary source of efficacy data.

NICE predominantly accepted synthetic control arms and propensity score-based analyses for decision-making but considered results to be uncertain due to the complexity of methods and potential bias.

Propensity-based ITC methods were more accepted than adjusting populations for MAICs, as they are more robust; unmatched variables lead to high uncertainty and potential for bias.

Comparisons with synthetic control or IPTW supported positive recommendations and were accepted by the ERG for use in economic base cases, in some cases.

Conclusions

While the potential benefits of innovative techniques like SCAs are widely discussed in the literature, our analysis suggests that standard indirect treatment comparison methods, such as matching-adjusted indirect comparisons, are still the dominant approach for generating comparative evidence in NICE submissions involving SATs, likely owing to a lack of availability of IPD for competitor trials.

Bias and confounders from observational and SAT data carry over into SCAs, reducing confidence in comparisons at HTA. In the future, AI/ML models may allow more precise matching of variables (eg, creating a "digital twin") and decrease uncertainty in indirect comparisons, which are still seen as inferior to RCTs at HTA.

More assessments and further experience of HTA body perspectives are required to determine if the theoretical benefits of synthetic approaches can be realized in practice.

References

1. Thorlund K, et al. ***Clin Epidemiol.*** 2020;12:457-467.
2. Selukar S, et al. ***Blood Rev.*** 2025;101324.
3. Patel D, et al. ***Value Health.*** 2021;24(8):1118-1125.
4. Phillippo D, et al. NICE Decision Support Unit Technical Support Document 18. 2016. http://www.nicedsu.org.uk. Accessed on 22 September 2025.

Poster presented at ISPOR Europe 2025, 9–12 November, Glasgow, UK