


ScienceDirect

Contents lists available at **sciencedirect.com**
Journal homepage: **www.elsevier.com/locate/jval**

Health Policy Analysis

Use of External Comparators for Health Technology Assessment Submissions Based on Single-Arm Trials

Dony Patel, MSc, PhD, Fiona Grimson, PhD, Elena Mihaylova, MSc, Peter Wagner, Joss Warren, Anke van Engen, MSc, Joseph Kim, PhD

ABSTRACT

Background: Single-arm trial (SAT) data is increasingly reviewed for drug approvals by regulators and Health Technology Assessment (HTA) bodies. Supplementary data in the form of external comparators (ECs) can be used to provide clinical context to support these drug evaluations. In this study we characterized HTAs for SAT-based submissions, the use of supplementary EC data and outcomes from HTA review.

Methods: HTA Accelerator database was used to describe SAT-based HTA submissions with decisions (2011-2019).

Results: A total of 433 SAT-based HTA submissions were identified between 2011 and 2019 with a 13-fold increase during this period. Around 65%(283/433) were in oncology or hem-oncology. Around 52%(226/433) of submissions contained some type of EC data, including prior clinical trials (24%, 104) and real-world data (RWD) (20%, 87), but 40%(175) contained no EC data. The overall acceptance rate for SAT-based submissions was 48% and with RWD EC data acceptance was 59%. In the latest 5-year period (2015-2019), use of RWD ECs increased 22% as a proportion of submissions per year, whereas, prior trial ECs decreased (−14%) and use of no EC remained stable (−2%). Between 2015 to 2017 and 2018 to 2019, acceptance rate for RWD ECs increased by 20% (41% in 2015-2017 to 61% in 2018-2019) whereas prior trial EC use decreased by 10% and no EC submissions decreased 16%. Of 226 submissions using ECs, only 29%(66) used an adjusted indirect treatment comparison method.

Conclusions: SAT-based submissions to HTA bodies are rapidly evolving in terms of composition and acceptance. Types of EC and methodological approach used are important determinants of positive outcomes.

Keywords: external comparators, health technology assessment, historical comparators, real-world data, synthetic comparators.

VALUE HEALTH. 2021; 24(8):1118–1125

Introduction

Randomized controlled trials (RCTs) are generally considered the gold standard evidence for marketing authorization by regulatory and Health Technology Assessment (HTA) bodies[1-3]. However, new drug development is increasingly focused on rare diseases and highly targeted patient populations, where traditional RCTs, controlled against placebo or standard of care, can present ethical and practical challenges.[1,2,4,5] Additionally, earlier phase trials may establish beyond clinical doubt that patients would benefit from the new treatment. In these circumstances non-RCT designs such as single-arm or multi-arm-uncontrolled studies, switchover studies or real-world studies[6-8] can be used.

Regulatory bodies have recognized the need for faster access to treatments in under-served patient populations. According to a recent review of European Medicines Agency (EMA) and United States Food and Drug Administration (FDA) submission,[8] most non-RCT development programs contained single-arm trial (SAT) designs. The EMA issued 795 and the FDA 774 SAT-based drug approvals between 1999 and 2014, with 44 and 60 respectively based solely on SAT evidence.[9,10] In an analysis of 253 United States FDA approvals issued between 2015 to 2017, 18% of the products had a pivotal trial with no comparator.[11]

This absence of internal controls in the evidence package is a challenge to HTA bodies, which generally prefer evidence from RCTs to determine the magnitude of benefit of a new product over standard of care.[3,12] When there is no direct comparison of products conducted within RCTs, multiple RCTs can be combined, and 'anchored' indirect treatment comparisons can be made using network meta-analysis methods.[13] When RCTs are not available, methods for 'unanchored' meta-analysis can still be used, such as matching-adjusted indirect comparisons (MAIC) and simulated treatment comparisons (STC).[14-16] When patient level data is available, more direct comparison methods can be applied such as

propensity score matching or Bayesian methods to help reduce selection bias and differences between comparator cohorts.[17-19] Some formal HTA guidance is available on these approaches, and their limitations from National Institute for Health and Care Excellence's (NICE) decision support unit.[20,21] In general, unanchored comparisons are less well established than anchored methods, and additional care is needed on interpretation.[12,22,23]

External comparator (EC) data (also referred to as "external controls," "historical comparators," or "synthetic controls") collected from patient cohorts outside of the clinical study, or "index study" is being used more frequently to provide contextual information on a product's clinical efficacy, safety, and cost-effectiveness. Here patient selection criteria from the index study protocol are applied to the external data source (to the extent possible) so that the comparator population looks as similar as possible to the index study population.[12,24] This type of supplementary dataset can be used for naive (unadjusted comparison) or adjusted indirect treatment comparison using STC or MAIC approaches.

To describe the landscape for HTA submissions on the basis of SAT data; the use of ECs in this context and their impact on HTA outcomes, we used IQVIA's HTA Accelerator platform to evaluate HTA reports from 100 HTA agencies across 40 countries published since 2011.

Methods

Database Overview

We performed a retrospective descriptive analysis using the IQVIA HTA Accelerator (www.iqvia.com/landing/hta-accelerator), which contains over 33 000 HTA publications covering 40 countries and 100 HTA bodies published since 2011. Data in HTA Accelerator comes from submissions, which are tracked in local language by market access experts who identify, translate and summarize new reports, and update existing reports when new information is available. Each report is captured and curated by analysts into a framework of 250 available data elements that includes indication, clinical evidence and comparators, economic analysis, HTA body critique, and recommendations. Internal workflows separate the data entry from quality review, with review being performed by senior team members. Automated logic checks are in place to further enhance database quality.

HTA outcomes were categorized as "positive" (including positive with restrictions), "negative," "no recommendation," or "multiple recommendations" (see Appendix Table 1 in Supplemental Materials found at https://doi.org/10.1016/j.jval.2021.01.015 for outcomes definitions). Positive and restricted positive recommendations were grouped for the purposes of outcomes analysis. One exception was for Haute Autorité de Santé (HAS) submissions in France where Amélioration du Service Médical Rendu (ASMR) V ratings were classified as negative recommendation (see Appendix Table 2 in Supplemental Materials found at https://doi.org/10.1016/j.jval.2021.01.015 for ASMR definitions). This rating indicates no added product benefit, and while products will still be reimbursed, subsequent price negotiations are affected. In Germany, Gemeinsamer Bundesausschuss (G-BA) no added benefit is automatically classified as a negative recommendation.

Datapoints were captured in addition to the standard HTA Accelerator dataset for this study. We categorized the type of EC as: (1) real-world data (RWD), with sub-categories as follows: registry, database study, chart review, and undefined RWD study; (2) prior trial; (3) RWD and prior trial; (4) unclear (including early access program data or expert opinions only); and (5) no EC. We considered all 5 categories mutually exclusive for this analysis but recognize that there are overlaps and that to fully describe the RWD study design, further details would be necessary. This was considered outside the scope of this analysis. Category 2 (prior trial), refers to cases where manufacturers employ either the active or control arm data from prior RCTs to compare with SAT data.

For each comparator data type, the method used for indirect treatment comparison (ITC) was classified as 'naive' (defined as side-by-side comparison with or without a statistical test) or 'adjusted' (when statistical adjustments were made to the match comparator population baseline characteristics - eg, MAIC, propensity scoring, or STC), where possible. We identified reports where these methods were used by searching the text for related terms and reviewing the context within reports.

HTA Submission Selection

HTA submissions were selected where a SAT was identified as the main type of evidence evaluated and any submissions with RCT evidence type were excluded. The data collection period was January 2011 to December 2019 based on the HTA decision publication date and only submissions where the HTA decision was available were selected. HTAs of first launch (original); label extensions as well as resubmissions, as they are known to contain new evidence, were included. Other types of submissions (renewals, change in formulation submissions, or unspecified type) were excluded from this analysis as they typically do not include any new data (see Appendix Table 3 in Supplemental Materials found at https://doi.org/10.1016/j.jval.2021.01.015 for a flow-chart describing case selection). HTA bodies that do not publish their appraisals would not be captured in this dataset. We found 433 HTA submissions fitting the study criteria from 21 countries (Appendix Table 4 in Supplemental Materials found at https://doi.org/10.1016/j.jval.2021.01.015).

We then created a reference set of HTA submissions as a broad comparison dataset, which contained evidence packages on the basis of all study-types including RCTs and non-RCT, with matching submission type and countries (Appendix Table 3 in Supplemental Materials found at https://doi.org/10.1016/j.jval.2021.01.015).

Data Analyses

Descriptive statistics were presented for overall SAT submissions per year and by therapy area. Submission characteristics were presented by type of supplementary or EC data and by HTA outcomes. Trends in HTA outcomes were analysed over the latest 5-year period, 2015 to 2019 by comparing decisions in between time period 2015 to 2017 and 2018 to 2019.

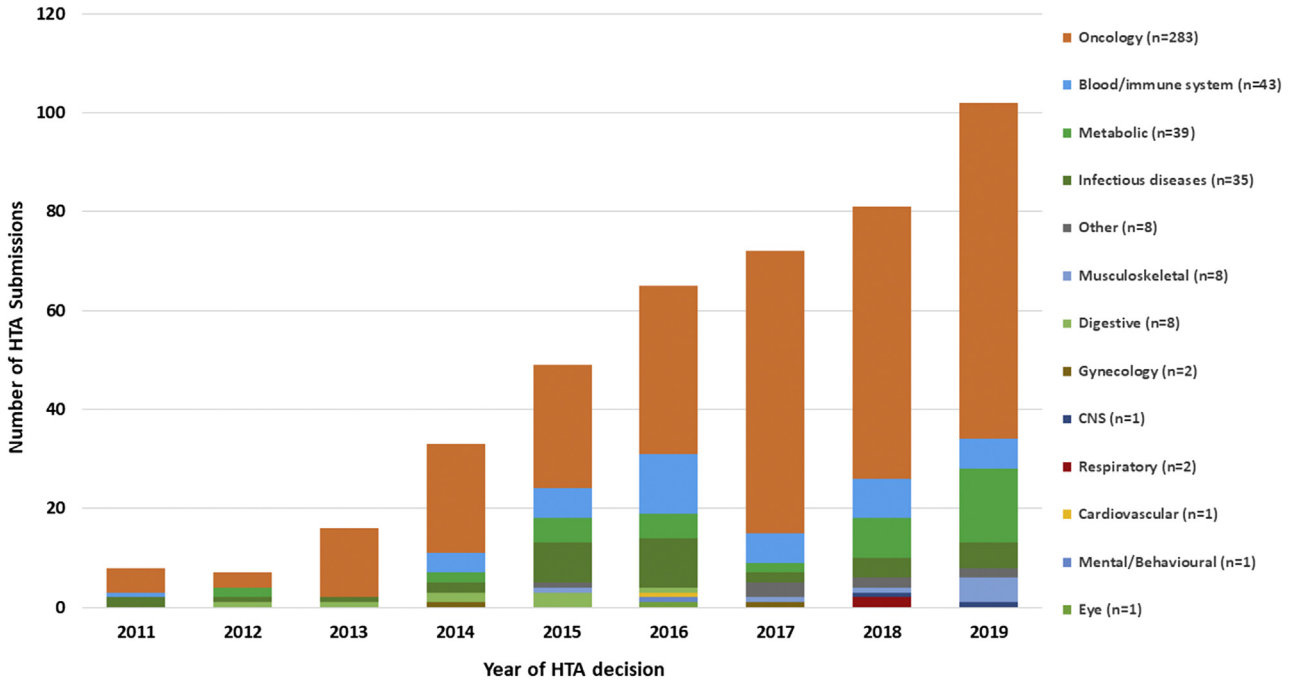
This analysis was further broken down by HTA for the Top 5 HTA bodies by submission volume covering pan-Canadian Oncology Drug Review (pCODR) or Canadian Agency for Drugs and Technologies in Health (CADTH), Frances' HAS, United Kingdom's NICE, Germany's G-BA and Australia's Pharmaceutical Benefits Advisory Committee (PBAC). Descriptive analysis was provided on the proportion of naive and adjusted statistical methods used for comparison by the EC category.

Results

Dataset Overview

We identified and reviewed 433 SAT-based HTA submissions (original, label extension or resubmissions) with HTA decisions available. This represented only 5% of the total submission volume between 2011 and 2019, but see a significant year-on-year increase from 8 to 102 over this period (Fig. 1). These 433

Figure 1. Single arm trial submissions to HTA bodies; Globally, 2011-2019.



submissions represented 166 individual drug/indication combinations with 65%(283/433) of submissions being for oncology or hemato-oncology indications. The next most frequent therapy areas were immune and/or blood disorders, endocrine and/or metabolic and infectious diseases (10%, 9% and 8%, respectively). 90(54%) of products/indication combinations were only submitted to one HTA body in a single country, 44(27%) recorded 2 to 4 different HTA submissions and 32(19%) recorded over 5 submissions.

Of the 40 countries covered in the HTA Accelerator platform, SAT-based HTA submissions were identified in 21 different countries with the top 5 being Canada (17%), France (16%), United Kingdom (15%), Germany (12%) and Australia (8.5%) (see Supplementary Table 4 in Supplemental Materials found at https://doi.org/10.1016/j.jval.2021.01.015 for full list).

This set of SAT-based submissions represented 275 (65%) original submissions, 126 (29%) for label extensions, and 32 (7%) resubmissions. Note this includes cases of multiple submissions to the same HTA body including an initial submission, then further resubmissions, possibly with additional datasets after feedback from the original submission.

Orphan drug designation and regulator conditional approval status was available for United States FDA and EMA countries, and accounts for 310 of the 433 SAT-based HTA submissions were identified. We found that 44% (137) included orphan drugs, with 82% (113) being for oncology indications. Across all oncology HTA submissions in these countries, 60% had orphan drug status. In addition, 33% (101/310) submissions were for drugs given a regulatory conditional approval and again nearly all (93%, 94) were for oncology indications. Notably, 50% (94/187) of all the SAT-based HTA submissions in oncology indications were conditionally approved by EMA or FDA. Non-orphan drug and conditional approval status were related, with 54% (74/137) of orphan drugs receiving the conditional approval of the regulator and only 16% (27/173) of non-orphan drugs receiving the conditional approval.

Types of ECs used in SAT Packages

Overall, 52.2% (226/433) of SAT submissions contained some type of supplementary dataset that can be designated as an EC. Based on information available, ECs were most often from prior clinical trials (24.0%, 104), then RWD (20.0%, 87), or both (8.1%, 35) (Table 1). We find that 40.4% (175) of submissions had no EC data. Finally, 7.4% (32) of submission packages contained supplementary evidence that was unclear, originating from expanded access programs, or based on expert opinion. Where we found ECs based on RWD only, details on the type of RWD source were determined to the possible extent. Of the 87 RWD ECs, 23(26.4%) were found to be database studies, 24 (27.5%) were registries, 3 (3.4%) were chart reviews, and the remaining 37 (42.5%) were of unclear design.

Trends in use of ECs were assessed over a 5-year period from 2015 to 2019. We see a 22% increase in the use of RWD-based ECs in this period (8%-30% - as a proportion of total SAT submissions per year). In contrast, use of prior clinical trials as ECs decreased 14% (from 33% to 17%). Overall, the use of ECs (RWD, prior clinical trials, or both) increased marginally (+6%) and the proportion of submissions with no EC remained relatively consistent (–2%) over the 5-year period.

HTA Outcomes Based on the Type of EC

Among the 433 HTA submissions on the basis of SAT data, 48%(208) received positive recommendations, including 33% (68/208) positive with restrictions. HTA submissions which included EC data received a positive recommendation in 52% (118/226) of the cases compared with 43% (75/175) with no EC.

Table 1. External comparator use and outcomes for single-arm trial based HTA submissions.

Source of comparator data	Recommendation/outcome (% of row total)				Total
	Multiple	Negative*	No Rec	Positive	
External comparator					
Real world data *(including registries, database or chart review studies)*	-	26 (30%)	10 (11%)	51 (59%)	87
Prior clinical trial(s)	-	30 (29%)	23 (22%)	51 (49%)	104
Prior clinical trial(s) and real world data	-	12 (34%)	7 (20%)	16 (46%)	35
Unclear *(including expanded access and expert opinion)*	1 (3%)	13 (41%)	3 (9%)	15 (47%)	32
None	3 (2%)	73 (42%)	24 (14%)	75 (43%)	175
Grand Total	4 (1%)	154 (36%)	67 (15%)	208 (48%)	433
Reference: All Submissions	720 (0.2%)	2741 (33%)	60 (0.7%)	5559 (66%)	8380

*HAS ASMR V classed as a negative outcome.

When assessing acceptance rates by type of EC, submissions which contained ECs based on RWD had most positive recommendations (59%, 50/87) followed by submissions with prior trial ECs (49%, 51/104). As a broad reference baseline, the HTA acceptance rate was 66%(5559/8380) for all HTA submissions, including RCT-based submissions for the same set of countries covering the same time period. Additionally, we found rates of acceptance of SAT-based submissions between categories of ECs were broadly consistent across the 3 major disease areas by volume (oncology, blood and/or immune, and endocrine and/or metabolic).

Trends in HTA outcomes were analyzed over the latest 5-year period, 2015 to 2019 by comparing decisions between time period 2015 to 2017 and 2018 to 2019 (Appendix Table 5 in Supplemental Materials found at https://doi.org/10.1016/j.jval.2021.01.015). Acceptance rate for RWD ECs increased +20% (41% in 2015-2017 to 61% in 2018-2019) whereas acceptance rate decreased for prior trial by -10% (54%-44%) and no EC by -15.5% (49.4%-33.9%).

Outcomes for SAT HTA Submissions for Key HTA Bodies

To understand the outcomes at individual HTA level, we reviewed comparator types and outcomes for the following key decision-making bodies: HAS, France (69); PBAC, Australia (37); G-BA, Germany (32); NICE, England (24); and pCODR/CADTH, Canada (33).

Among these HTA bodies, only NICE showed substantially more positive than negative recommendations (83%, 20 vs 17%,4) (Fig. 2). This is in line with the overall acceptance rate for NICE which is 85%. Importantly 80% (16/20) of NICE positive decisions for SAT-based submissions had restricted access with 11 cases controlled through the cancer drug fund. In other HTA bodies, positive recommendation rates were only marginally higher than negative recommendations. The proportion of positive acceptances with restrictions were lower for G-BA (21%, 4/19), pCODR/CADTH (23%, 5/22), and PBAC (38%, 6/16). In France, HAS received most SAT-based submission of all HTA bodies, as it reviews the broadest range of products of all HTAs along with pCODR/CADTH. For HAS, we find 36% positive outcomes for SAT-based submissions and 61% were negative. However, of these negative decisions, 76% (32/42) were classified as ASMR V, which allows for reimbursement negotiations where the product is considered to add no benefit over available current treatments.

When assessing outcomes by the type of comparator, we find that RWD and prior clinical trial ECs were more often found as part of positive submissions than negative ones in Canada, United Kingdom, Germany, and France. Only in Australia, we found higher number of RWD comparators in submissions with negative outcomes. However, interpretation should be cautioned as sample size is small when analyzing individual HTA bodies.

Orphan drugs have a different threshold of evidence requirements compared with nonorphan drugs. In Germany, orphan drugs (with annual turnover threshold below €50 million) are assumed to have an added benefit.[25] We analyzed the HTA outcomes for NICE, HAS, and G-BA for which EMA orphan designation was available (125 submissions). A total of 49 SAT-based HTA submissions to these bodies had orphan designation and 37 (76%) received a positive decision. In comparison, out of 76 nonorphan submissions, only 27 (37%) were positive. G-BA gave positive decisions for 87% (12/14) of the orphan drug submissions and for NICE, 93%(13/14) were positive. In France, HAS gave 57% (12/21) of orphan submissions a positive decision, while 7 were classified as negative decisions with ASMR V. Therefore, statistical tests for significance of outcome differences would need to account for orphan status as well as other observed confounding factors.

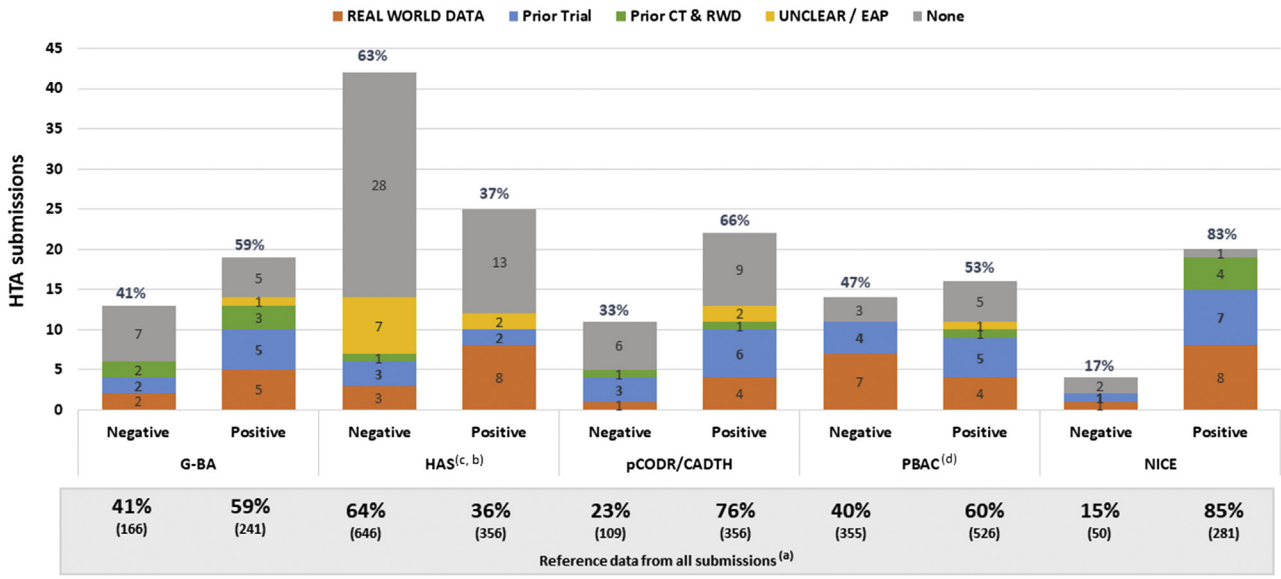
Methodological Approaches for ECs

EC studies can vary in approach, not only by data source but also by method of comparison (Table 2). Here we present analysis on whether the ITC used a statistical adjustment or a naive (nonadjusted) comparison approach.

In the 226 submissions using ECs, only 29% (66) of EC data was used for an adjusted ITC. This represents only 16% (66/433) of all SAT submissions. 61%(137) were presumed to use ECs in a naive comparison context. A prior review of regulatory submissions found that only 9% (4/43) of non-RCT submissions used an adjusted ITC.[23] In this analysis, the first case of adjusted ITC was not until 2015 (3, 7% of 2015 submissions) increasing to 29% (30) in 2019. Over the same period, use of naive ITC reduced as a proportion of all submissions between 2015 (22, 48%) and 2019 (22, 22%). Prior clinical trials were the single largest source for ECs and had the highest proportion of adjusted ITC (31%, 32/104). When all RWD EC sources are combined, 25% (22/87) were used for adjusted ITC.

Assessing outcomes according to comparison methodology, no clear differences in overall outcomes were found. Overall, adjusted ITC submissions had 47% (32/68) positive rate and naive ITC submissions had 53% (77/144) positive rate. We did see a difference when analyzing trends in acceptance over the latest 5-year period by comparing time periods 2015 to 2017 and 2018 to 2019 (Appendix Table 5 in Supplemental Materials found at https://doi.org/10.1016/j.jval.2021.01.015). We find adjusted ITC

Figure 2. External comparator submissions and outcomes for top 5 HTA bodies by volume, 2011-2019.



(a) % positive and negative outcomes for all HTA submissions within the period January 2011 to December 2019 , (b) HAS ASMR V rating (no added benefit) classified as negative for this analysis, (c)HAS SAT submissions also included 1 case with no recommendation and 1 case with multiple recommendations , (d)PBAC SAT submissions also included 7 cases with multiple recommendations explained as either PBAC awaiting the regulatory decision as this can run in parallel or waiting for an update on price for resubmission

positive decisions increased by 14% (36% in 2015-2017 vs 51% in 2018-2019), whereas naive ITC decisions decreased by 8% (56% vs 48%).

Discussion

To enable access to medicines in situations where RCTs are not ethical or feasible, an understanding of the landscape and outcomes for SAT-based submissions is necessary. We found the use of SAT-based submissions comprises only 5% of total submissions to HTA bodies; in line with a prior study of 3 HTA key bodies (NICE, CADTH, and Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen [IQWiG]) in which 4% to 6% of submissions were solely based on non-comparative clinical trial evidence.[22] However, we found a significant yearly increase over this period, making this category of submission packages important to monitor, and this study may support HTA bodies/payers in developing guidelines and policies to handle these submissions. To the best of the author's knowledge, this is the first review of SAT-based submissions covering all published HTA appraisals and the use of ECs.

General Trends in EC use

Most SAT-based submissions are in oncology/hemato-oncology indications, as is expected, based on the current industry trend for novel targeted therapies and cell and gene therapy approaches for rare cancers and biomarker-targeted sub-populations.[26] However, we were surprised by the finding that only about half of the SAT-

Table 2. Method of comparison by type of external comparator (EC).

Source of comparator data*, n (%)	Method of comparison			Total
	Adjusted ITC	Naïve ITC	Unclear/No EC	
Prior clinical trial	32 (31%)	62 (60%)	10 (10%)	104
Undefined RWD study	8 (22%)	23 (62%)	6 (16%)	37
Prior CT & RWD study	12 (34%)	20 (57%)	3 (9%)	35
Registry	4 (17%)	17 (71%)	3 (13%)	24
Database study	7 (30%)	15 (65%)	1 (4%)	23
Chart review	3 (100%)			3
Grand Total	66 (29%)	137 (61%)	23 (10%)	226

RWD indicates Real World Data; ITC, indirect treatment comparison; CT, clinical trial.
*Categories considered as mutually exclusive for purposes of this analysis.

based submissions contained some type of supplementary EC data, with only a marginal increase since 2015. However, this is higher in proportion than that observed for regulatory submissions, on the basis of a previous studies of EMA and FDA submissions, which reported that 37% of regulatory submissions contained an EC between 2010 to 2015.[8,10,27,28] We expect HTA bodies to increasingly demand ECs to assess incremental benefit of new drugs and our research supports this trend. We find increasing use and acceptance of RWD ECs and decrease in positive outcomes for submissions with no EC between 2015 to 2019. The reasons cited for negative HTA decisions commonly mention missing clinical comparison through a common comparator, lack of an appropriate comparator, and uncertainly in the new drugs' clinical benefit.

The statistical methods for external comparison are an ongoing area of research. In this analysis, we did not see clear differences in outcomes between adjusted and naive (nonadjusted) ITC methods overall. However, in the latest 5-year period, we find increasing use of adjusted EC comparisons and a trend toward increasing acceptance. This suggests that HTA bodies are beginning to show a preference for population adjusted unanchored ITC methods such as STC and MAIC over nonadjusted methods.[20,22] Moreover, we see HTA commentaries for naive EC comparison criticizing a lack of adjustment for confounders and uncertainty owing to potential bias in results.

Impact of ECs on HTA Decision-Making

Our study assessed outcomes across 21 HTA bodies and we acknowledge that this covers the broad array of agencies that have important differences in terms of maturity, familiarity with nontraditional evidence packages, the extent to which they make information publicly available, decision-making remit, and other country-specific nuances.[29,30] In this context, we find that adding RWD ECs resulted in a higher acceptance rate than observed when prior trial ECs or no EC data at all were used, which explains the observed increase in the use of RWD and decrease in the use of prior clinical trial data between 2014 and 2019. The increasing availability of RWD and gradual acceptance by HTA and regulatory bodies has been extensively covered by other commentators.[23,31-34] In relation to SAT-based submissions, RWD ECs may be considered more appropriate as they allow more relevant comparator cohorts to be constructed and enable the collection of more contemporary standard of care treatments than prior clinical trials data. However, we suggest that manufacturers assess other important factors when designing an EC as there is variation between HTA bodies in acceptance of RWD. In addition, there are important advantages of historical trial data to consider such as the availability of a wide range of endpoints for comparison. Consequently, in our review of HTA commentaries, we found it challenging to find clear reasons for higher acceptability for RWD ECs over other types of EC datasets and this point warrants further study.

For a deeper understanding of country HTA specific factors for decision making, we evaluated report summaries from key HTA bodies. For NICE, we see highest acceptance rate ($>$ 80%), which reflects their acceptance for nontraditional evidence package, including RWD and familiarity with EC approaches.[14,20] However, most NICE SAT-based submissions were for oncology indications, which were given restricted access via the cancer drugs fund. Thus, allowing NICE to carry out further evaluation to manage uncertainties in clinical benefit. In the France, HAS classified a high proportion of outcomes as negative with the important caveat that most negative decisions were ASMR V ratings. This means HAS could not find any added benefit for the product over existing treatments, subsequently leading to less favorable pricing. Our study shows, manufacturers that included RWD-based ECs in their submissions mostly received higher ASMR ratings (Fig. 2). HAS also issued ASMR V outcomes with an additional request to provide further study data either from pending clinical trials or RWD studies, and, thus, this outcome could be considered another type of conditional approval. These findings suggest a deeper review of HTA commentaries is justified to give further insights regarding use of conditional approvals and re-evaluations for SAT-based submissions.

We know that HTA decisions are influenced by factors such as unmet need and orphan designation, which makes it a challenge when trying to correlate the HTA outcome to the strength of the evidence package. Our results show orphan status is an important determinant in acceptability of SAT-based submissions, as 58% of orphan drug submissions receive positive HTA outcomes compared with 38% of nonorphan drug submissions. HTA bodies have a higher evidentiary threshold for non-orphan drug submissions, and, therefore, even greater reliance on robust EC data. Our data show non-orphan submissions with RWD ECs are more like to receive positive outcomes (71%, 5/7), whereas, submissions using prior trial ECs or no comparators have mostly negative HTA outcomes. We are however cautious with this conclusion because of the low sample size for this calculation.

Limitations

The findings from our study are intended to provide a description of trends in use of EC data in support HTA decisions, and not to provide a causal link between the use of RWD or EC data and the probability of a successful submission. This review was conducted based on reports from HTA bodies captured systematically within HTA Accelerator; however, there are important limitations to consider. First, we are limited to reports published publicly by HTA bodies. While many HTA bodies are mandated to publish all appraisals, several HTA bodies do not publish appraisals, such as Japan and Italy. Second, where appraisals are published, the level of information varies. In particular, the level of detail on the evidence package related to the SAT and supplementary data including the EC was limited in some published appraisals. Therefore, the findings of this study may have been influenced by publication bias. Additionally, while our review aims to be exhaustive, there is a chance that we did not identify all the HTA submissions on the basis of pivotal SAT data. This may be the case in situations where the pivotal trial uses a single-arm design, but there are also supportive trial data based on a randomized design, which mean these submissions would not be included in the current review. Additionally, we did not attempt to examine all decision drivers for drug acceptance, owing to the large volume of commentary to review. Additionally, communication between the manufacturer and the HTA body, which may help to understand the rationale for the evidence package may be missing in the published reports. Finally, our review intentionally focused on SAT packages. This is because we understand this to the predominant design for nonrandomized trials and because SATs can be clearly identifiable within the HTA Accelerator platform. However, we acknowledge that other nonrandomized study designs are used for regulator and HTA submissions.

Conclusion

Our analysis describes the complexities in assembling a strong evidence package for SAT-based submissions for HTA review and the current variably in HTA outcomes. Consequently, it illustrates the importance of having a well-considered strategy when relying on an SAT, robust data collection, and a well-executed analytical approach

to the indirect treatment comparison. There is little formal guidance or established best practice for appraisal of HTA submissions based on SAT evidence, which makes the construction of comparator cohorts challenging. Careful evaluation of past HTA decisions and early input from HTA stakeholders are critical to designing an EC study that can support correct contextualization of SATs.

Supplemental Materials

Supplementary data associated with this article can be found in the online version at https://doi.org/10.1016/j.jval.2021.01.015.

Article and Author Information

Accepted for Publication: January 24, 2021

Published Online: June 8, 2021

doi: https://doi.org/10.1016/j.jval.2021.01.015

Author Affiliations: Real World Solutions (Patel) and EMEA Centre of Excellence for Retrospective Studies (Grimson, Kim), IQVIA, London, England, UK; IQVIA R&D Solutions, Sofia, Bulgaria (Mihaylova); Consulting Services, IQVIA, Munich, Germany (Wagner); Real World Solutions, IQVIA, Durham NC, USA (Warren); Real World Solutions, IQVIA, Netherlands (Engen); Faculty of Epidemiology and Population, London School of Hygiene and Tropical Medicine, London, England, UK (Kim); Department of Health Policy, London School of Economics and Political Science, London, England, UK (Kim).

Correspondence: Dony Patel, MSc, PhD, IQVIA Real World Solutions, 210 Pentonville Road, London, England, United Kingdom N1 9YJ. Email: dony.patel@iqvia.com

Author Contributions: *Concept and design*: Patel, Mihaylova, Warren, van Engen, Kim
Acquisition of data: Patel, Mihaylova, Wagner
Analysis and interpretation of data: Patel, Grimson, Mihaylova, Wagner, van Engen
Drafting of the manuscript: Patel, Grimson, Mihaylova, Wagner, Warren, Kim
Critical revision of the paper for important intellectual content: Patel, Grimson, Wagner, Warren, van Engen, Kim
Statistical analysis: Patel, Grimson
Supervision: Patel, van Engen, Kim

Conflict of Interest Disclosures: Drs Patel, Grimson, Mihaylova, Wagner, Warren, van Engen, and Kim are employed by IQVIA.

Funding/Support: This study was financially supported by IQVIA in the form of salaries for all authors. The study was conducted independent of additional funding.

Role of the Funder: The funder had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

REFERENCES

1. ICH harmonised tripartite guideline: choice of control group and related issues in clinical trials E10. Paper presented at: International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use; July 20, 2000. https://database.ich.org/sites/default/files/E10_Guideline.pdf. Accessed May 26, 2020.
2. U.S. Food and Drug Administration. Rare diseases: natural history studies for drug development guidance for industry. https://www.fda.gov/regulatory-information/search-fda-guidance-documents/rare-diseases-natural-history-studies-drug-development. Accessed May 26, 2019.
3. Moloney R, Mohr P, Hawe E, Shah K, Garau M, Towse A. Payer perspectives on future acceptability of comparative effectiveness and relative effectiveness research. *Int J Technol Assess Health Care*. 2015;31(1-2):90–98.
4. Attwood MM, Rask-Andersen M, Schiöth HB. Orphan drugs and their impact on pharmaceutical development: [published correction appears in *Trends Pharmacol Sci*. 2018;39(12):1077]. *Trends Pharmacol Sci*. 2018;39(6):525–535.
5. Pontes C, Fontanet JM, Vives R, et al. Evidence supporting regulatory-decision making on orphan medicinal products authorisation in Europe: methodological uncertainties. *Orphanet J Rare Dis*. 2018;13(1):206.
6. Elsevier. Glossary of methodologic terms. https://www.elsevier.com/__data/promis_misc/apmrglossary.pdf. Accessed April 1, 2020.
7. U.S. National Library of Medicine. Glossary of common site terms. https://clinicaltrials.gov/ct2/about-studies/glossary. Accessed April 1, 2020.
8. 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.
9. Beaver JA, Howie LJ, Pelosof L, et al. A 25-year experience of US food and drug administration accelerated approval of malignant hematology and oncology drugs and biologics: a review. *JAMA Oncol*. 2018;4(6):849–856.
10. Hatswell AJ, Baio G, Berlin JA, Irs A, Freemantle N. Regulatory approval of pharmaceuticals without a randomised controlled study: analysis of EMA and FDA approvals 1999-2014. *BMJ Open*. 2016;6(6):e011666.
11. Gores M. Seal of approval: accelerating regulatory success with RWE. https://www.iqvia.com/library/white-papers/seal-of-approval-accelerating-regulatory-success-with-rwe. Accessed May 26, 2020.
12. Gray CM, Grimson F, Layton D, Pocock S, Kim J. A framework for methodological choice and evidence assessment for studies using external comparators from real-world data. *Drug Saf*. 2020;43(7):623–633.
13. Hoaglin DC, Hawkins N, Jansen JP, et al. Conducting indirect-treatment-comparison and network-meta-analysis studies: report of the ISPOR task force on indirect treatment comparisons good research practices: part 2. *Value Health*. 2011;14(4):429–437.
14. Phillippo DM, Ades AE, Dias S, Palmer S, Abrams KR, Welton NJ. Methods for population-adjusted indirect comparisons in health technology appraisal. *Med Decis Making*. 2018;38(2):200–211.
15. Bucher HC, Guyatt GH, Griffith LE, Walter SD. The results of direct and indirect treatment comparisons in meta-analysis of randomized controlled trials. *J Clin Epidemiol*. 1997;50(6):683–691.
16. Jiang Y, Ni W. Performance of unanchored matching-adjusted indirect comparison (MAIC) for the evidence synthesis of single-arm trials with time-to-event outcomes. *BMC Med Res Methodol*. 2020;20(1):241.
17. Lin J, Gamalo-Siebers M, Tiwari R. Propensity score matched augmented controls in randomized clinical trials: a case study. *Pharm Stat*. 2018;17(5):629–647.
18. Chen WC, Wang C, Li H, et al. Propensity score-integrated composite likelihood approach for augmenting the control arm of a randomized controlled trial by incorporating real-world data. *J Biopharm Stat*. 2020;30(3):508–520.
19. Ghadessi M, Tang R, Zhou J, et al. A roadmap to using historical controls in clinical trials - by Drug Information Association Adaptive Design Scientific Working Group (DIA-ADSWG). *Orphanet J Rare Dis*. 2020;15(1):69.
20. Phillippo DM, Ades AE, Dias S, Palmer S, Abrams KR, Welton NJ. NICE DSU technical support document 18: methods for population-adjusted indirect comparisons in submisions to NICE. http://nicedsu.org.uk/wp-content/uploads/2017/05/Population-adjustment-TSD-FINAL.pdf. Accessed May 26, 2020.
21. NICE decision support unit. Observational data TSD. http://nicedsu.org.uk/technical-support-documents/observational-data-tsd/. Accessed October 28, 2020.
22. Griffiths EA, Macaulay R, Vadlamudi NK, Uddin J, Samuels ER. The role of noncomparative evidence in health technology assessment decisions. *Value Health*. 2017;20(10):1245–1251.
23. Pinto A, Naci H, Neez E, Mossialos E. Association between the use of surrogate measures in pivotal trials and health technology assessment decisions: a retrospective analysis of NICE and CADTH reviews of cancer drugs. *Value Health*. 2020;23(3):319–327.
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.
25. Verfahrensordnung G-BA §12.2 based on the version effective as of October 13, 2020, G.B. (G-BA). Editor.
26. IQVIA. Launch Excellence V: surviving and thriving when launching in an increasingly specialised world. https://www.iqvia.com/library/white-papers/launch-excellence-v. Accessed May 26, 2020.
27. Makady A, de Boer A, Hillege H, Klungel O, Goettsch W. (on behalf of GetReal Work Package 1). What is real-world data? A review of definitions based on literature and stakeholder interviews. *Value Health*. 2017;20(7):858–865.
28. Sasinowski FJ, Panico EB, Valentine JE. Quantum of effectiveness evidence in FDA's approval of orphan drugs: update, July 2010 to June 2014. *Ther Innov Regul Sci*. 2015;49(5):680–697.
29. Allen N, Pichler F, Wang T, Patel S, Salek S. Development of archetypes for non-ranking classification and comparison of European National Health Technology Assessment systems. *Health Policy*. 2013;113(3):305–312.
30. Allen N, Liberti L, Walker SR, Salek S. A comparison of reimbursement recommendations by European HTA agencies: is there opportunity for further alignment? *Front Pharmacol*. 2017;8:384.
31. Government of Canada. A strategy to optimize the use of real-world evidence across the medical device life cycle in Canada. https://www.canada.ca/en/health-canada/corporate/transparency/regulatory-transparency-and-openness/improving-review-drugs-devices/real-world-evidence-medical-device-strategy.html. Accessed April 26, 2020.
32. U.S. Food and Drug Administration. Submitting documents using real-world data and real-world evidence to FDA for drugs and biologics guidance for

industry. https://www.fda.gov/regulatory-information/search-fda-guidance-documents/submitting-documents-using-real-world-data-and-real-world-evidence-fda-drugs-and-biologics-guidance. Accessed April 26, 2020.
33. Bolislis WR, Fay M, Kühler TC. Use of real-world data for new drug applications and line extensions. *Clin Ther*. 2020;42(5):926–938.
34. European Medicines Agency. HMA-EMA joint big data task force phase II report: 'evolving data-driven regulation'. https://www.ema.europa.eu/en/documents/other/hma-ema-joint-big-data-taskforce-phase-ii-report-evolving-data-driven-regulation_en.pdf. Accessed April 26, 2020.