BMJ Open

Real-world data from expanded access programs in health technology assessments: a review of NICE technology appraisals

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2021-052186
Article Type:	Original research
Date Submitted by the Author:	07-Apr-2021
Complete List of Authors:	Polak, Tobias; Erasmus University Rotterdam Institute of Health Policy and Management, Health Technology Assessment; Erasmus MC, Department of Biostatistics Cucchi, David; Amsterdam UMC Locatie VUmc, Department of Haematology van Rosmalen, Joost; Erasmus MC, Biostatistics Uyl-de Groot, Carin; Erasmus Universiteit Rotterdam, Erasmus School of Health Policy & Management
Keywords:	GENERAL MEDICINE (see Internal Medicine), HEALTH ECONOMICS, Health policy < HEALTH SERVICES ADMINISTRATION & MANAGEMENT

I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

Real-world data from expanded access programs in health technology assessments: a review of NICE technology appraisals

Tobias B. Polak M.Sc.[123], David G.J. Cucchi M.D.[4], Joost van Rosmalen Ph.D.[2], and Carin A. Uyl - De Groot Ph.D.[1]

[1] Erasmus School of Health Policy & Management, Erasmus University Rotterdam, Rotterdam, The Netherlands

[2] Department of Biostatistics, Erasmus Medical Center, Rotterdam, The Netherlands

[3] RWD Department, myTomorrows, Amsterdam, The Netherlands

[4] Department of Hematology, Amsterdam UMC, Vrije Universiteit Amsterdam, Cancer Center Amsterdam, Amsterdam, The Netherlands

Running head

How NICE uses RWD from expanded access programs

Corresponding author

Tobias B. Polak

Department of Biostatistics, Na28

Doctor Molewaterplein 40

3015 GD Rotterdam, the Netherlands

t.polak@erasmusmc.nl, +31612626025

Keywords

Health technology assessment; real-world data; expanded access; NICE; technology appraisals

Word count

3633

Abstract

Introduction

Expanded access (EA) offers patients that are ineligible for clinical trials or registered treatment options access to investigational therapies. EA programs are increasingly used to collect real-world data (RWD) in a pre-approval setting. RWD and data from clinical trials are used by the National Institute for Health and Care Excellence (NICE) to conduct cost-effectiveness analyses for novel technologies.

Objective

To quantify and characterize the usage of EA data in NICE technology appraisals.

Methods

Cross-sectional study of NICE appraisals (2010-2020). We automatically downloaded and screened all available appraisal documentation on NICE website (over 8500 documents), searching for EA-related terms. Two reviewers independently labeled the EA usage for disease area, and whether it was used to inform safety, efficacy, and/or resource use.

Results

In 54.2% (206/380 appraisals) at least one reference to EA was made. 20.8% (79/380) of the TAs used EA data to inform safety (n=43), efficacy (n=47) and/or resource use (n=51). The number of TAs that utilize EA data increased over time (p=0.009), and EA data utilization was disproportionally distributed across disease areas (p=0.001).

Conclusion

NICE uses EA data in over one in five appraisals. In synthesis with evidence from well-controlled trials, data from EA programs increasingly informs cost-effectiveness modeling.

Article Summary

- Expanded access programs are progressively used by patients and physicians to speed access to investigational medicine, and by regulators to gain insight in real-world usage of novel therapies.
- This study is the first to assess how health technology assessments rely on data from expanded access programs.
- We evaluated all NICE appraisal documentation from 2010 to 2021.

Strengths and limitations

- We demonstrate that combining automated and manual screening efficiently facilitates health policy analyses.
- Our search was limited to health technology appraisals performed by NICE.

Introduction

Novel drug therapies are important drivers of increased health care spending. In the United Kingdom (UK), the National Institute for Health and Care Excellence (NICE) conducts technology appraisals (TAs) to evaluate technologies (e.g. drugs, medical devices) on cost-effectiveness and to determine their impact on health care budgets[1]. These evaluations are conducted using a variety of data sources, such as randomized controlled trials (RCTs) or observational studies [2,3]. In this research, we explore the use of data in NICE TAs from another source: expanded access programs.

A positive appraisal determination from NICE forms the main pathway for novel pharmaceutical technologies to access the National Health Service (NHS) and become available for patients across the UK. The health technology assessment (HTA) usually starts with the submission of evidence on clinical effectiveness and costs by the pharmaceutical company. The submission is scrutinized by an independent Evidence Review Group (ERG), which critically reviews the manufacturer's submission and performs additional exploratory analyses of cost-effectiveness [1,4–6].

Patients, patient advocacy groups, and physicians working within the NHS also contribute to NICE's appraisals. The resulting qualitative input is considered in the formal analyses conducted by the manufacturer and the ERG. The entire evidence is assessed by NICE's Appraisal Committee and forms the basis of their appraisal determination[6]. More detailed information on NICE's processes can be found on their guidance website (https://www.nice.org.uk/about/what-we-do/our-programmes/nice-guidance)

HTA bodies are particularly keen to know how technologies will use resources, yield benefit, and attribute risks in the real-world patient population for which treatment will potentially be reimbursed [7]. Real-World Data (RWD) are 'information on health care that is derived from multiple sources outside typical clinical research settings', such as electronic health records, claims and billing databases, or patient registries [8]. RWD is typically generated after a drug comes to market (post-approval). At the time of reimbursement decision however, most of the available data stems from clinical trials (pre-approval). Noteworthy, payers may use (real-world) data from patients that have been treated outside of clinical trial settings, but prior to marketing authorization [1–3]. These patients can receive treatment via expanded access programs.

Expanded access (EA) is a pathway for patients who suffer from life-threatening conditions, who cannot enter clinical trials, and have exhausted all approved treatment options, to access investigational medicine. It is also known as 'compassionate use', 'early access' or 'non-trial pre-approval access' [9]. The primary intent of EA programs is to provide patients and physicians in dire need with potential treatment options outside of clinical trials. Secondary, such programs may be used to collect RWD. It offers a potential opportunity to collect real world data in a pre-approval setting [10–12].

Data from EA programs may be used for various purposes in the appraisal process, e.g. to inform formal safety or efficacy analyses, to inform resource cost in real-world settings, to estimate the patient population, or to serve as qualitative input by patients or physicians that have treatment experience within an EA program. Although its use by regulators has recently been researched[10], the use of EA data by payers or HTA bodies remains unquantified. Understanding its use in TAs clarifies the value of these data for payers, pharmaceutical industry, physicians and patients, and is relevant for cost-effectiveness decision making and evaluation of HTA policy. Therefore, we here investigate the role EA data play in payers decision making by reviewing all appraisals presented to NICE between 2010 and 2021.

Methods

Documents relating to all TAs conducted are provided on the NICE website. We investigated TAs published between 01-01-2010 and 01-01-2021. Terminated, withdrawn, or replaced appraisals were removed as documentation was unavailable. A schematic overview of our workflow is provided in **Figure 1**.

We wrote a computer script (i.e. a web scraper [10]) to automatically list and download all documentation (e.g. manufacturer submissions, ERG report, final appraisal determination) available through NICE's website. Subsequently, the script extracted the text from these documents and automatically screened whether the text contained 'expanded access terms', like 'Compassionate Use', 'Expanded Access' 'Early Access', etc, as well as all possible spellings thereof. When at least one of these 'expanded access terms' were present, two authors (T.B.P. and D.G.J.C.) independently, manually, reviewed the context of term.

We primarily labeled the data usage with one or more of the following categories:

1. Safety: EA data were used to evaluate the safety profile

2. Efficacy: EA data were used to evaluate the efficacy profile

3. Resource use: EA data were used to inform cost parameters

4. Trivial: EA data were not used or trivially mentioned in the appraisal

Patient and physicians also share their treatment experience. As the impact of these accounts is harder to quantify, we did not include them in our main analysis but secondarily labeled:

1. Treatment experience: When patients or physicians cited experience within the EA program.

Discordance was resolved by discussion between the two reviewers. To give the reader a sense of these different types of usage, examples are provided in the Results section. Furthermore, TAs were classified as single technology appraisal (STA), multiple technology appraisal (MTA), or highly specialized technology (HST). All TAs were categorized according to their area of disease.

Statistics

Spearman rank correlation test was used to detect time trends in the yearly number of appraisals using EA data. We performed a Pearson chi-square test to assess differences in distribution of disease areas between appraisals that did include or did not include EA data. For all significance testing, we set the 2-sided significance level at 0.05.

Results

We screened all 496 TAs conducted between 01-01-2010 and 01-01-2021. This ranged from TA185 to TA667 and from HST1 to HST13. N=116 appraisals were excluded (for details, see **Figure 1**). The remaining 380 appraisals had 8925 documents that were downloaded and screened.

In 54.2% (206 of 380 appraisals) at least one reference to EA was made. In total, 79 out of 380 (20.8%) of the TAs used EA data to inform safety (n=43), efficacy (n=47) or resource use (n=51). As a single TA could have multiple labels, there is overlap between safety, efficacy and resource use. This is depicted in **Figure 2A**. Additionally, in 54 appraisals (14.5%) the EA program was cited by patients or physicians as treatment experience.

There is a significant increase over time in the absolute use of EA data by payers (ρ = 0.74 and p = 0.009; **Figure 2B**). However, the number of TAs also increases over time. There is no evidence of a significant increase in use of EA over time relative to the total number of appraisals conducted (ρ = 0.32 and p = 0.332).

Significant differences (χ^2= 38.8, p= 0.001) exist in the disease areas that did versus those that did not include EA data. Oncology and hematology together account for 66% of the appraisals with EA data, whereas they only make up 50% of the entire fraction of appraisals. On the other hand, disease areas such as cardiology, gastroenterology, endocrinology, dermatology, rheumatology and ophthalmology jointly make up 24.5% of all appraisals, whereas they merely account for 2.6% of the appraisals that included EA data. These results can be found in **Table 1.**

Examples

To give the reader a better sense of the main labels 'safety, efficacy, resource use' as well as the secondary 'treatment experience' label, we here provide illustrative examples from the TAs that were supported by EA data.

Safety

Safety data from EA programs are often described rather qualitatively, supporting results from clinical trials. For example, in the appraisal of gefitinib for the first-line treatment of locally advanced or metastatic non-small-cell lung cancer, the appraisal committee noted:

> *"The favorable safety profile of gefitinib demonstrated in the phase III studies is consistent with that observed in everyday settings. In addition to the data from clinical trials, the Early Access Program for gefitinib in Caucasian patients indicated that gefitinib is well tolerated by patients with advanced or metastatic NSCLC. The majority of ADRs associated with gefitinib are mild in nature and those most commonly reported are grade 1/2 diarrhoea and skin reactions".*
>
> Manufacturer submission, Safety and tolerability, TA192

Alternatively, safety signals from EA programs can be quantitatively incorporated in modeling disutilities that jointly determine cost-effectiveness. When evaluating ocrelizumab for treating relapsing–remitting multiple sclerosis, the committee noted that an important safety signal from the compassionate use program is lacking from the current analysis:

> *"The committee heard that there has been the 1 case of PML* (progressive multifocal leukoencephalopathy, red.) *following treatment with ocrelizumab in the compassionate-use programme in Germany, (...). It concluded that the economic model should have included a risk of PML for ocrelizumab".*
>
> Appraisal consultation, Adverse events in the economic model, TA533

Efficacy

Efficacy data from EA programs can also be used, together with data from clinical trials, to determine overall efficacy of the technology appraised. In the evaluation of lutetium (177Lu) oxodotreotide for treating irresectable or metastatic neuroendocrine tumors, response rates were obtained from the 'Erasmus study'. The Erasmus study was a compassionate use program conducted at the Erasmus Medical Centre. The data from this program are summarized as:

> *"In a single centre non-controlled phase I/II open-label study* (The Erasmus study, red.)*, conducted in 810 Dutch patients with different somatostatin receptor positive tumour types, the objective response rate (ORR) for the full analysis set (FAS) population with GEP-NETs and bronchial NETs (360 patients) was 44% (95% confidence interval [CI] 38% - 49%)."*
>
> Manufacturer submission, Executive summary, TA539

Payers are likely to assess benefit in terms of quality-adjusted life years (QALYs), as their main decision to recommend or not recommend a product for reimbursement depends (among other things) on the willingness-to-pay for an incremental year in perfect health. In the evaluation of cabazitaxel for hormone-relapsed metastatic prostate cancer treated with docetaxel, the EA program was used to gather quality of life data not collected during the routine clinical development:

> *"The company did not collect data on health-related quality of life in TROPIC* (the RCT, red.)*, so it took utility values from the UK Early Access Programme (EAP) for cabazitaxel. The programme measured the health-related quality of life (using the EQ-5D) of men who had been treated with cabazitaxel after docetaxel.(...)"*
>
> *"(...) One hundred and twelve patients participated in the UK EAP at 12 UK Cancer Centres. All had mCRPC with disease progression during or after docetaxel and were similar in baseline patient characteristics to the*

population in TROPIC. (...) Safety assessments were performed prior to each cycle and HRQL recorded at alternate cycles using the EQ-5D-3L questionnaire and visual analogue scale (VAS).”

Committee papers, Health-related quality of life, TA391

Resource use

EA data can also be used to inform other parameters in cost-effectiveness modeling. Such models are often based on Markov chains, that describe the state of the disease that patients are in at a given time point. These models require cost per state and transition probabilities between states. Registries, or other real-world data sources, are frequently used to estimate such data. In the appraisal of sofosbuvir-velpatasvir-voxilaprevir for treating chronic hepatitis C, transition probabilities from decompensated liver cirrhosis to death are modeled via a Beta-distribution and the input parameters are provided from the EA program:

"Variable: From decompensated cirrhosis to death

Distribution and parameters: Beta; α=46.5; β=147.2

Source: EAP data (expanded access program, red.)"

Manufacturer submission, Sensitivity analyses, TA507

A different, direct resource use example is given in the evaluation of ipilimumab for previously treated irresectable malignant melanoma. The dosing of ipilimumab is weight-dependent. Hence, to estimate the number of vials needed for treatment of UK patients, an estimate of the (UK) patient population weight is required. This weight is calculated via:

"Patient level analysis of the weight of UK clinical trial patients in MDX010-20 (n=55), and the weight of UK patients in the ipilimumab compassionate use program (n=258), from these weights, the mean number of vials required (assuming no vial sharing) is calculated."

“Results from these analyses showed that the dose of ipilimumab given per patient per induction has a large impact on the ICER with the minimum dose given in the trial and compassionate use programme (3 x 50 mg) resulting in an ICER of £38,387 per QALY gained and the maximum dose (2 x 200 mg) given resulting in an ICER of £88,788 per QALY gained.”

Manufacturer submission, Intervention and comparators costs, TA268

Treatment experience

NHS professionals share their opinions and experience on the technology appraised in expert committee meetings. In the appraisal of patisiran for treating hereditary transthyretin amyloidosis, the Head of the National Amyloidosis Centre (NAC) is asked 'how data on real-world experience in this condition compare with clinical trial data?'. His response is:

> *"The experience of my colleagues at the NAC treating patients through compassionate access (over one year) and Early Access to Medicine Schemes has been extremely favourable. Remarkable clinically significant improvements of well-being and function have occurred in a majority of cases, including regaining the ability to walk unaided."*
>
> Clinical expert statement, HST10

Patients, caregivers, or patient group representatives are also provided the opportunity to share their experience with the appraised treatment. The assessment of nusinersen for treating muscular atrophy sparked comments from parents with children that suffer from this disease:

> *"My son is currently receiving Spinraza at Gosh for type 1c SMA. He was lucky enough to be included into the expanded access program for a select group of children. Since receiving his treatment we have watched the transformation of a seriously weakening child to a thriving boy who has gained significant progress in his motor function and health, we are continually amazed by his progress. He starts preschool in the coming weeks, an achievement we never thought possible. (...)"*
>
> Patient/caregiver stakeholder comment, TA588

Discussion

In this review, we combined automated documentation searches with double, independent manual review to screen NICE documentation on the usage of EA data. We have found NICE to frequently review data from EA programs to support their decision making: 20.8% of the TAs used EA data to evaluate safety, efficacy/effectiveness, or resource

use of the appraised technology. The use of data from EA programs appears to increase over the years. Additionally, patients and physicians share their treatment experience with the EA program in 14.2% of the appraisals.

The disease areas of the appraisals that included EA data differed significantly from the overall distribution of disease areas from all appraisals investigated between 2010 and 2021. Oncology and hematology account for the lion's share (66%) of EA data usage, yet account for half (50%) of all TAs conducted. Although 'the life-threatening or seriously debilitating' prerequisite for EA is often present in hemato-oncologic malignancies, cardiac or ophthalmologic illnesses can also be severely limiting[13,14]. Cardiology and ophthalmology account for 8.4% of all TAs, but none (0%) of these programs used EA data (or even mentioned it). Drug developers in these areas may be less familiar with collecting and using EA data, or cardiologists and ophthalmologists may be less acquainted with EA than hemato-oncologists.

Compared to regulatory submissions to the EMA and the FDA, submissions to NICE more frequently include EA data. The EMA and FDA used EA data to support efficacy in 49 regulatory approvals over 25 years ($\mp$ 2 annually)[10]. In this work, we find that NICE used EA to inform cost-effectiveness in 76 over 11 years ($\mp$ 7.2 annually). One reason for this may be that payors have a higher uptake of RWD in their decision making. Second, regulators primarily consider safety and efficacy, whereas HTA bodies look at safety, efficacy, and at resource use. Furthermore, they also assess comparative effectiveness rather than efficacy. Investigating resource use and comparative effectiveness simply adds degrees of freedom to NICE's decision making process and one of those components can be jointly determined by EA data. Modeling cost and comparative effectiveness by definition necessitates a variety of input parameters, every one of them potentially coming from different sources.

Whether using EA data for payer decision making is wise, depends wholly on the decision at hand. The instances in which the FDA and the EMA assessed efficacy mainly based on EA data, are scarce, and characterized by (i) a high unmet medical need (ii) a rare disease population and (iii) large treatment effects[10]. These three prerequisites do not often hold up. Additionally, we witnessed twice (TA391, TA491) that HRQoL were not gathered during the conventional clinical trials but were captured in the EA program. Although data from EA program can bridge an evidence gap, HRQoL data should simply have been collected during all stages of clinical development. For safety, the primary evidence assessment comes from RCTs, although the use of registries, post-approval safety studies, or pharmacovigilance during EA, is widely recognized. The textbook example would be to detect infrequently occurring

adverse events from observational datasets. Indeed, we identified such an example in TA533, where the compassionate use program led to the identification of a rare but serious adverse event. Overall, the evidence for assessing safety and efficacy should primary come from RCTs and can be synthesized with RWD or other non-randomized sources, such as EA programs.

Including EA can have several advantages, as it can increase sample size, inform additional parameters - such as HRQoL - or aid to estimate effects for patients that were excluded from the trial, but were included in the EA program. Such patients are generally older and frailer [7,15,16], and thus collecting data in these populations helps to extrapolate results on safety and efficacy found in RCTs. Estimates of resource use parameters that are derived from clinical trials, such as adherence, monitoring, or the number of hospital visits, can even be more distinct from real-world settings. Therefore, EA data can play a useful role in informing resource use parameters. Furthermore, modeling resource use requires a vast amount of input parameters, such as transition rates for disease states, or disease incidence, that are collected over longer periods of time. This is infeasible to gather from clinical trials and thus patient or population registries or EA programs may inform decision making.

The regulatory status of data collection during EA programs is a matter of debate [10,11,17–20]. In Europe, individual Member States regulate EA programs[21]. Different countries may issue conflicting statements that can be at cross with EMA decision making[10]. This also resonates in appraisals. For example, we read in the appraisal of cemiplimab for treating metastatic or locally advanced cutaneous squamous cell carcinoma:

> *"While formal data collection is not permitted from a regulatory standpoint, the safety of cemiplimab at the flat 350mg dose in a real-world setting will be monitored."*
>
> Manufacturer submission, Safety overview, TA592

This begs the questions who decides what formal and informal data collection is and whether all examples put forth in this paper where impermissible for regulators. Regardless of regulatory requirements, it can be a source of frustration when EA data are not available, as one Advisory Group (AG) noted:

> *"The lack of any efficacy data from the compassionate use program is particularly disappointing,"*
>
> AG response to company comments, AG conclusions, TA535

Although the primary intent of EA programs is treatment provision and not to conduct research, it seems awkward to treat patients with investigational medicine and not to collect data to inform safety and efficacy. Furthermore, it is difficult to precisely determine where treatment-intent ends and research-intent starts. The changing nature of EA programs from sole treatment-intent to treatment-intent with data collection is a current topic of debate among bioethicists [12,18,22]. We stress that data collection during EA should be light-weight and must not disproportionally burden patient and physicians – hence, a smart design should facilitate data to be collected [12]. If so, EA programs can be the first source of RWD to inform HTA evaluations gathered in a pre-approval setting – this makes EA data different from general RWD sources (e.g. electronic health records or claims and billing data), as the latter will only start generating evidence once the drug has been approved.

Limitations and future research

Our work has several limitations. First, we only reviewed TAs from one HTA body: NICE. Formally, NICE's decisions are only valid within their UK jurisdiction, but informally they lead the way for other European HTA bodies - either via setting an example or via reference pricing. We have chosen NICE for our review as they have the longest history of HTA assessment and ample documentation publicly available. For other HTA bodies, results may be different. Future research should confirm whether our results uphold for other HTA bodies. Preliminary findings presented at a conference concluded that using EA data gathered within French compassionate use programs had a positive impact on reimbursement discussions [23]. Second, we may have missed use-cases of EA data in payer submission as companies or reviewers may have used other terms to indicate EA programs (or failed to have done so). Our automated algorithm facilitates high throughput of document screening in health policy analysis, but it may have missed cases that would have been identified in manual evaluation. Therefore, our estimates should be interpreted as lower bound of EA use in NICE appraisals. Lastly, we were unable to exactly quantify the added value of EA data. As we lack a counterfactual, we do not know what would have happened without the inclusion of EA data. Additionally, it is difficult to measure the 'impact' of EA data, as it is not always clear how these data have exactly been used. Therefore we have provided the reader with both high-level quantitative statistics and with qualitative samples from our data set.

Conclusion

NICE uses EA data in over one in five (20.8%) of their appraisals and this number appears to increase over time. In general, adding data from EA can yield more real-world information. Especially to estimate the resource use, pre-

approval EA data can play a vital role informing post-approval real-world usage. In synthesis with evidence from well-controlled trials, data collected from EA programs meaningfully informs NICE decision making.

Author contributions

T.B.P., D.G.J.C, J.v.R, and C.A.U. - d.G. contributed to the concept and design. T.B.P. and D.G.J.C. acquired and analysed the data. T.B.P. Drafted the manuscript. D.G.J.C, J.v.R, and C.A.U. - d.G. revised the manuscript.

Funding

This research was supported via an unrestricted grant from HealthHolland: EMCLSH20012. HealthHolland is a funding vehicle for the Dutch Ministry of Economic Affairs and Climate Policy that addresses the Dutch Life Sciences & Health sector.

Conflict of interest

C.A.U. - d.G. has received unrestricted grants from Boehringer Ingelheim, Astellas, Celgene, Sanofi, Janssen-Cilag, Bayer, Amgen, Genzyme, Merck, Glycostem Therapeutics, AstraZeneca, Roche, and Merck. T.B.P. works part-time at myTomorrows and holds stock in myTomorrows. J.v.R and D.G.J.C. declare no conflict of interest.

Data availability

Raw data and code are available on request via the corresponding author.

Tables

Table 1: *Technology appraisals that did ('Yes') or did not ('No') include Expanded Access (EA) data to support the profile of safety, efficacy and/or resource use, classified on disease area.*

	Included EA data			
	No[1]	Yes[1]	**Total**[1]	p-value[2]
Disease area				0.001
Benign hematology	5 (1.7%)	3 (3.8%)	8 (2.1%)	
Cardiology	14 (4.7%)	0 (0%)	14 (3.7%)	
Dermatology	12 (4.0%)	1 (1.3%)	13 (3.4%)	
Endocrinology	12 (4.0%)	0 (0%)	12 (3.2%)	
Gastroenterology	13 (4.3%)	0 (0%)	13 (3.4%)	
Hematology	35 (12%)	20 (25%)	55 (14%)	
Internal medicine	23 (7.6%)	9 (11%)	32 (8.4%)	
Neurology	14 (4.7%)	6 (7.6%)	20 (5.3%)	
Oncology	106 (35%)	32 (41%)	138 (36%)	
Ophthalmology	18 (6.0%)	0 (0%)	18 (4.7%)	
Psychiatry	3 (1.0%)	1 (1.3%)	4 (1.1%)	
Pulmonology	6 (2.0%)	4 (5.1%)	10 (2.6%)	
Rheumatology	22 (7.3%)	1 (1.3%)	23 (6.1%)	
Surgery	4 (1.3%)	1 (1.3%)	5 (1.3%)	

	Included EA data			
	No[1]	Yes[1]	**Total**[1]	p-value[2]
Urology	1 (0.3%)	1 (1.3%)	2 (0.5%)	
Vascular medicine	13 (4.3%)	0 (0%)	13 (3.4%)	
Total	301 (79%)	79 (21%)	380 (100%)	

[1] *n (%)*

[2] *Pearson chi-square test*

Table legend

Table 1: Technology appraisals that did ('Yes') or did not ('No') include Expanded Access (EA) data to support the profile of safety, efficacy and/or resource use, classified on disease area.

References

1. The National Institute for Health and Care. Guide to the methods of technology appraisal. *Nice*. 2018;(April 2013):1-93.

2. Bell H, Wailoo AJ, Hernandez M, et al. The use of real world data for the estimation of treatment effects in NICE decision making. 2016;2016(June):1-61.

3. Faria R, Hernández Alava M, Manca A, Wailoo AJ. The use of observational data to inform estimates of treatment effectiveness in technology appraisal: Methods for comparative individual patient data. *NICE Decis Support Unit*. 2015;(0):0-85.

4. National Institute for Health and Clinical Excellence (NICE). *Guide to the Single Technology Appraisal Process*.; 2009.

5. National Institute for Health and Clinical Excellence (NICE). *Guide to the Multiple Technology Appraisal Process*.; 2009.

6. Dias S, Welton NJ, Sutton AJ, Ades AE. Evidence synthesis for decision making 1: Introduction. *Med Decis Mak*. 2013;33(5):597-606. doi:10.1177/0272989X13487604

7. Blommestein H. The added value of real-world evidence. Published online 2016.

8. Sherman RE, Anderson SA, Dal Pan GJ, et al. Real-World Evidence — What Is It and What Can It Tell Us? *N Engl J Med*. 2016;375(23):2293-2297. doi:10.1056/NEJMsb1609216

9. Kimberly LL, Beuttler MM, Shen M, Caplan AL, Bateman-House A. Pre-approval Access Terminology. *Ther Innov Regul Sci*. 2017;51(4):494-500. doi:10.1177/2168479017696267

10. Polak TB, van Rosmalen J, Uyl - de Groot CA. Expanded access as a source of real-world data: An overview of FDA and EMA approvals. *Br J Clin Pharmacol*. Published online March 22, 2020. doi:10.1111/bcp.14284

11. Polak TB, van Rosmalen J, Uyl - de Groot CA. Response to Open Peer Commentary " Making It Count : Extracting Real World Data from Compassionate Use and Expanded Access Programs " Response to Open Peer Commentary " Making It Count : Extracting Real World Data from Compassionate Use and Expanded

Acces. *Am J Bioeth*. 2020;20(11):W4-W5. doi:10.1080/15265161.2020.1820113

12. Rozenberg O, Greenbaum D. Making It Count: Extracting Real World Data from Compassionate Use and Expanded Access Programs. *Am J Bioeth*. 2020;20(7):89-92. doi:10.1080/15265161.2020.1779857

13. Taylor DJ, Hobby AE, Binns AM, Crabb DP. How does age-related macular degeneration affect real-world visual ability and quality of life? A systematic review. *BMJ Open*. 2016;6(12):e011504. doi:10.1136/bmjopen-2016-011504

14. Chandra A, Lewis EF, Claggett BL, et al. Effects of Sacubitril/Valsartan on Physical and Social Activity Limitations in Patients With Heart Failure: A Secondary Analysis of the PARADIGM-HF Trial. *JAMA Cardiol*. 2018;3(6):498-505. doi:10.1001/jamacardio.2018.0398

15. Westgeest HM, Uyl-de Groot CA, van Moorselaar RJA, et al. Differences in Trial and Real-world Populations in the Dutch Castration-resistant Prostate Cancer Registry. *Eur Urol Focus*. 2018;4(5):694-701. doi:10.1016/j.euf.2016.09.008

16. Munk NE, Knudsen JS, Pottegård A, Witte DR, Thomsen RW. Differences Between Randomized Clinical Trial Participants and Real-World Empagliflozin Users and the Changes in Their Glycated Hemoglobin Levels. *JAMA Netw Open*. 2020;3(2):e1920949. doi:10.1001/jamanetworkopen.2019.20949

17. Bunnik EM, Aarts N, van de Vathorst S. Little to lose and no other options: Ethical issues in efforts to facilitate expanded access to investigational drugs. *Health Policy*. 2018;122(9):977-983. doi:10.1016/j.healthpol.2018.06.005

18. Chapman CR, Moch KI, McFadyen A, et al. What compassionate use means for gene therapies. *Nat Biotechnol*. 2019;37(4):352-355. doi:10.1038/s41587-019-0081-7

19. Rawson K. Expanded Access Data Can Support Approval Decisions, US FDA Says. Accessed June 19, 2019. https://pink.pharmaintelligence.informa.com/PS124296/Expanded-Access-Data-Can-Support-Approval-Decisions-US-FDA-Says

20. S. Usdin. Beyond compassionate use: the case for using expanded access protocols to generate real world data. Accessed June 9, 2019. https://www.biocentury.com/biocentury/regulation/2017-09-29/case-using-

expanded-access-protocols-generate-real-world-data

21. Balasubramanian G, Morampudi S, Chhabra P, Gowda A, Zomorodi B. An overview of compassionate use programs in the european union member states. *Intractable Rare Dis Res*. 2016;5(4):244-254. doi:10.5582/irdr.2016.01054

22. Bunnik EM, Aarts N, van de Vathorst S. The changing landscape of expanded access to investigational drugs for patients with unmet medical needs: ethical implications. *J Pharm Policy Pract*. 2017;10(1):10. doi:10.1186/s40545-017-0100-3

23. Villemur L, Cohignac L, Jeanbat V, Bouee S. PNS266 Collecting EARLY Access Real World Evidence DATA for a Successful Reimbursement Appraisal. *Value Heal*. 2020;23:S685. doi:10.1016/j.jval.2020.08.1710

Figure legends

Figure 1. Screening and selection of technology appraisals from NICE. STA, single technology appraisal. MTA, multiple technology appraisals. HST, highly specialized technology. EA, expanded access. NICE, National Institute for Health and Care Excellence

Figure 2: Technology appraisals (TAs) using Expanded Access (EA) data to support safety, efficacy and/or resource use. A: Venn-diagram displaying the overlap of safety, efficacy, and/or resource use labeling of TAs. **B**: Bar chart of TAs published between 01-01-2010 and 01-01-2021 that did ('Yes') or did not ('No') include data EA programs to support safety, efficacy and/or resource use.

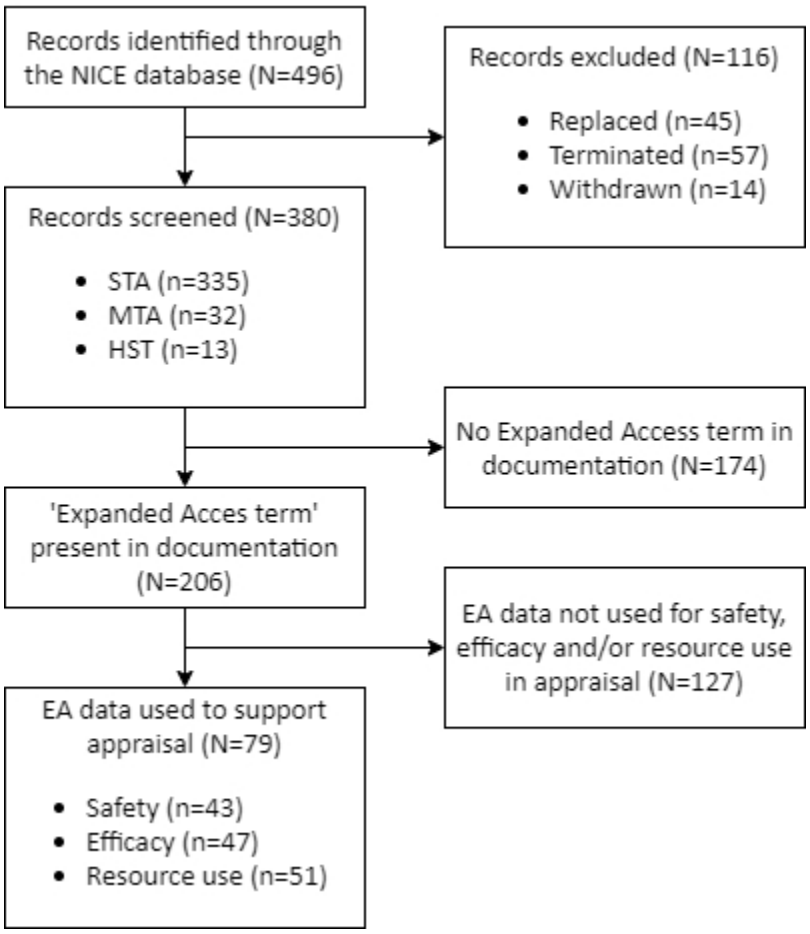


Figure 1. Screening and selection of technology appraisals from NICE. STA, single technology appraisal. MTA, multiple technology appraisals. HST, highly specialized technology. EA, expanded access. NICE, National Institute for Health and Care Excellence

141x164mm (72 x 72 DPI)

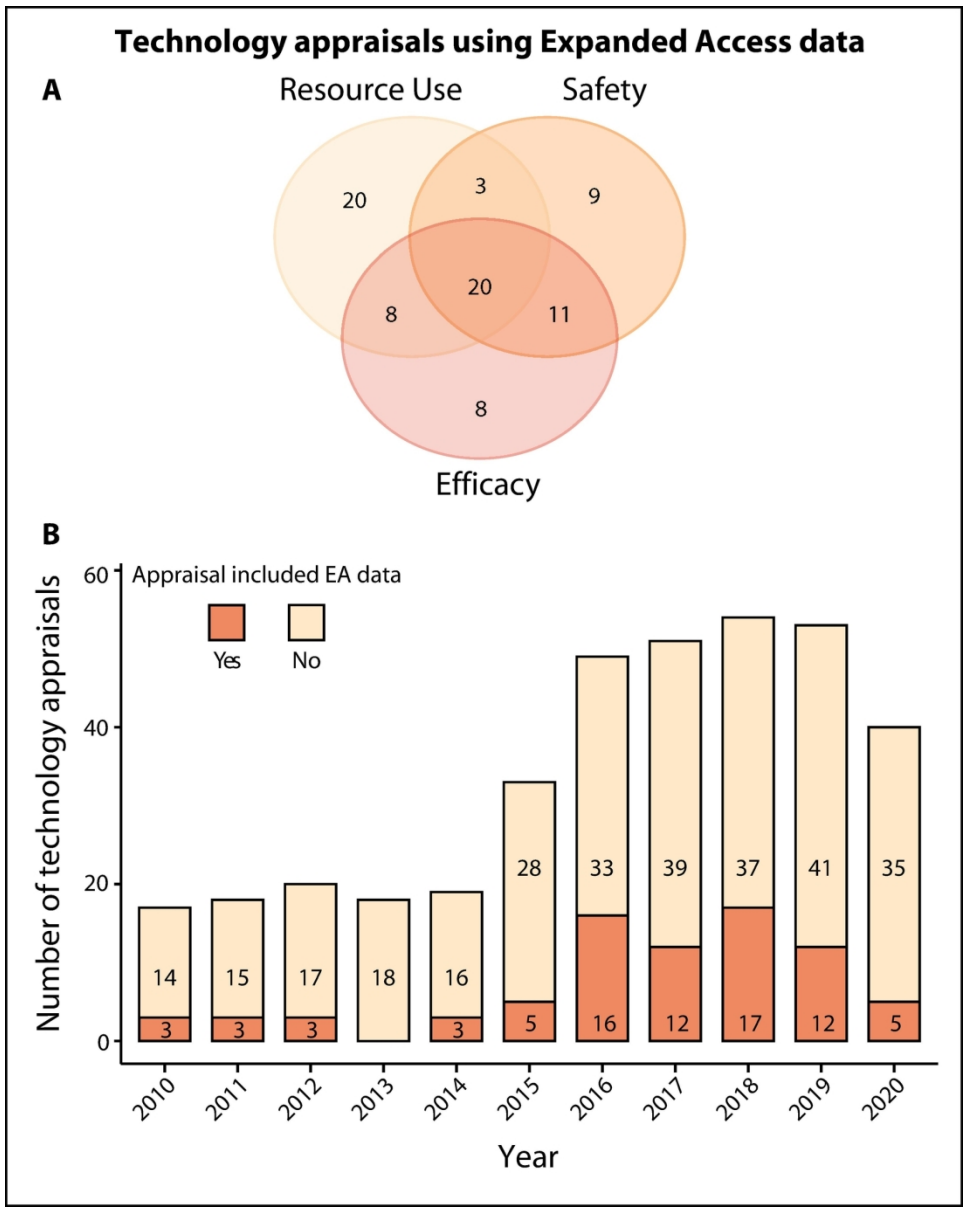


Figure 2: Technology appraisals (TAs) using Expanded Access (EA) data to support safety, efficacy and/or resource use. A: Venn-diagram displaying the overlap of safety, efficacy, and/or resource use labeling of TAs. B: Bar chart of TAs published between 01-01-2010 and 01-01-2021 that did ('Yes') or did not ('No') include data EA programs to support safety, efficacy and/or resource use.

153x191mm (300 x 300 DPI)

BMJ Open

Real-world data from expanded access programmes in health technology assessments: a review of NICE technology appraisals

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2021-052186.R1
Article Type:	Original research
Date Submitted by the Author:	08-Nov-2021
Complete List of Authors:	Polak, Tobias; Erasmus University Rotterdam Institute of Health Policy and Management, Health Technology Assessment; Erasmus MC, Department of Biostatistics Cucchi, David; Amsterdam UMC Locatie VUmc, Department of Haematology van Rosmalen, Joost; Erasmus MC, Biostatistics Uyl-de Groot, Carin; Erasmus Universiteit Rotterdam, Erasmus School of Health Policy & Management
Primary Subject Heading:	Health policy
Secondary Subject Heading:	Health economics
Keywords:	GENERAL MEDICINE (see Internal Medicine), HEALTH ECONOMICS, Health policy < HEALTH SERVICES ADMINISTRATION & MANAGEMENT

I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

Real-world data from expanded access programmes in health technology assessments: a review of NICE technology appraisals

Tobias B. Polak M.Sc.[1,2,3,4], David G.J. Cucchi M.D., Ph.D.[5], Joost van Rosmalen Ph.D.[2,3], and Carin A. Uyl - De Groot Ph.D.[1]

[1] Erasmus School of Health Policy & Management, Erasmus University Rotterdam, Rotterdam, The Netherlands

[2] Department of Biostatistics, Erasmus Medical Centre, Rotterdam, The Netherlands

[3] Department of Epidemiology, Erasmus Medical Centre, Rotterdam, The Netherlands

[4] RWD Department, myTomorrows, Amsterdam, The Netherlands

[5] Department of Haematology, Amsterdam UMC, Vrije Universiteit Amsterdam, Cancer Centre Amsterdam, Amsterdam, The Netherlands

Running head

Use of RWD from expanded access programmes by NICE.

Corresponding author

Tobias B. Polak

Department of Biostatistics, Na28

Doctor Molewaterplein 40

3015 GD Rotterdam, The Netherlands

t.polak@erasmusmc.nl, +31612626025

Keywords

Health technology assessment; real-world data; expanded access; NICE; technology appraisals

Word count

3969

Abstract

Objectives

To quantify and characterize the usage of expanded access (EA) data in National Institute for Health and Care Excellence (NICE) technology appraisals (TAs). EA offers patients that are ineligible for clinical trials or registered treatment options, access to investigational therapies. Although EA programmes are increasingly used to collect real-world data (RWD), it is unknown if and how these date are used in NICE health technology assessments.

Design

Cross-sectional study of NICE appraisals (2010-2020). We automatically downloaded and screened all available appraisal documentation on NICE website (over 8500 documents), searching for EA-related terms. Two reviewers independently labelled the EA usage by disease area, and whether it was used to inform safety, efficacy, and/or resource use. We qualitatively describe the 5 appraisals with the most occurrences of EA-related terms.

Primary outcome measure

Number of technology appraisals that used expanded access data to inform safety, efficacy and/or resource use analyses.

Results

In 54.2% (206/380 appraisals) at least one reference to EA was made. 21.1% (80/380) of the TAs used EA data to inform safety (n=43), efficacy (n=47) and/or resource use (n=52). The number of TAs that utilize EA data remained stable over time, and the extent of EA data utilisation varied by disease area (p=0.001).

Conclusion

NICE uses EA data in over one in five appraisals. In synthesis with evidence from well-controlled trials, data collected from EA programmes may meaningfully inform cost-effectiveness modelling.

Strengths and Limitations

- This study is the first to assess whether health technology assessments rely on data from expanded access programmes.
- Our search was limited to health technology appraisals performed by NICE between 2010 and 2020.
- Combining automated and manual screening can efficiently facilitate health policy analyses.

Introduction

Novel drug therapies are important drivers of increased health care spending. In the United Kingdom (UK), the National Institute for Health and Care Excellence (NICE) conducts technology appraisals (TAs) to evaluate cost-effectiveness of technologies (e.g. drugs, medical devices) and to determine their impact on health care budgets[1]. These evaluations are conducted using a variety of data sources, such as randomized controlled trials (RCTs) or observational studies [2,3]. In this research, we explore the use of data in NICE TAs from another source: expanded access programmes.

A positive appraisal determination from NICE forms the main pathway for novel pharmaceutical technologies to access the National Health Service (NHS) and become available for patients across the UK. The health technology assessment (HTA) usually starts with the submission of evidence on clinical effectiveness and costs by the pharmaceutical company. The submission is scrutinized by an independent Evidence Review Group (ERG), which critically reviews the manufacturer's submission and performs additional exploratory analyses of cost-effectiveness; in some cases the ERG even re-analyses clinical data[1,4–6].

Patients, patient advocacy groups, and physicians working within the NHS also contribute to NICE's appraisals. The resulting qualitative input is considered in the formal analyses conducted by the manufacturer and the ERG. The entire evidence is assessed by NICE's Appraisal Committee and forms the basis of their appraisal determination[6]. More detailed information on NICE's processes can be found on their guidance website (https://www.nice.org.uk/about/what-we-do/our-programmes/nice-guidance).

HTA bodies are particularly keen to know how technologies will use resources, yield benefit, and attribute risks in the real-world patient population for which treatment will potentially be reimbursed[7]. Real-World Data (RWD) are 'information on health care that is derived from multiple sources outside typical clinical research settings', such as electronic health records, claims and billing databases, or patient registries[8]. RWD is typically generated after a drug comes to market (post-approval). At the time of the reimbursement decision however, most of the available data stems from clinical trials (pre-approval). Noteworthy, payers may use (real-world) data from patients that have been treated outside of clinical trial settings, but prior to marketing authorization[1–3]. These patients can receive treatment via expanded access programmes.

Expanded access (EA) is a pathway to access investigational medicine for patients who suffer from life-threatening conditions, who cannot enter clinical trials, and have exhausted all approved treatment options. It is also known as 'compassionate use', 'early access' or 'non-trial pre-approval access'[9]. The primary intent of EA programmes is to provide patients and physicians in dire need with potential treatment options outside of clinical trials. Secondary, such programmes may potentially collect real world data in a regulatory pre-approval setting, but the generation and useability of evidence derived from these programmes remains a topic of debate.[10–16]

Data from EA programmes may be used for various purposes in the appraisal process, for example to inform formal safety or efficacy analyses, to inform resource use and associated costs in real-world settings, to estimate the size of the patient population, or to gain insights into the treatment experience from patients or physicians that participated in an EA programme. These data are increasingly accepted to support evidence of clinical efficacy by regulators, especially when collecting data in controlled settings is infeasible, such as in (ultra-)rare diseases, or is deemed unethical, in the case of extremely large treatment effects[10]. However, the use of EA data by payers or HTA bodies remains unquantified. Understanding the role of EA data in TAs may clarify the value of these data for payers, pharmaceutical industry, physicians and patients, and is relevant for cost-effectiveness decision making and evaluation of HTA policy. Therefore, we here investigate the usage of EA data in NICE decision making by reviewing all appraisals presented to NICE between 2010 and 2020.

Methods

Documents relating to all TAs conducted are provided on the NICE website. We investigated TAs published between 01-01-2010 and 01-01-2021. Terminated, withdrawn, or replaced appraisals were removed as documentation was unavailable. A schematic overview of our workflow is provided in **Figure 1**.

We wrote a computer script (i.e. a web scraper[10]) to automatically list and download all documentation (e.g. manufacturer submissions, ERG report, final appraisal determination) available through NICE's website. Subsequently, the script extracted the text from these documents and automatically screened whether the text contained 'expanded access (EA) terms', like 'Compassionate Use', 'Expanded Access' 'Early Access', etc., as well as all possible spellings thereof. A detailed protocol, including all search terms, are available in the **Supplementary Files A**. The data and code from the paper are available on the GitHub from the first author,

https://github.com/TobiasPolak. When at least one of these 'expanded access terms' were present, two authors (T.B.P. and D.G.J.C.) independently and manually, reviewed the context of the term.

We primarily labelled the data usage with one or more of the following categories:

1. Safety: EA data were used to evaluate the safety profile
2. Efficacy: EA data were used to evaluate the efficacy profile
3. Resource use: EA data were used to inform cost parameters
4. Trivial: EA data were not used or trivially mentioned in the appraisal

Patient and physicians also share their treatment experience. As the impact of these accounts is harder to quantify, we did not include them in our main analysis but secondarily labelled:

1. Treatment experience: When patients or physicians cited experience within the EA programme.

Discordance was resolved by discussion between the two reviewers. To give the reader a sense of these different types of usage, examples are provided in the Results section. Additionally, we provide a narrative summary of the 5 appraisals that contain the most occurrences of the search terms to illustrate the use of EA data qualitatively. Lastly, TAs were classified as single technology appraisal (STA), multiple technology appraisal (MTA), or highly specialized technology (HST). All TAs were categorized according to their area of disease.

Patient and Public involvement

No patients were involved during the planning and writing of this work; all data were derived from NICE technology appraisals.

Statistics

The Spearman rank correlation test was used to detect time trends in the yearly number of appraisals using EA data. We performed a Pearson chi-square test to assess whether the proportion of appraisals that included EA data differed by disease area. For all significance testing, we set the 2-sided significance level at 0.05.

Results

We screened all 496 TAs conducted between 01-01-2010 and 01-01-2021. This ranged from Technology Appraisal 185 (TA185) to TA667 and from Highly Specialized Technology 1 (HST1) to HST13. *N*=116 appraisals were excluded (for details, see **Figure 1**). The remaining 380 appraisals had 8925 documents that were downloaded and screened.

In 54.2% (206 of 380 appraisals) at least one reference to EA was made. In total, 80 out of 380 (21.1%) of the TAs used EA data to inform safety (n=43), efficacy (n=47) or resource use (n=52). As a single TA could have multiple labels, there is overlap between safety, efficacy and resource use. This is depicted in **Figure 2A**. Additionally, in 54 appraisals (14.5%) the EA programme was cited by patients or physicians as treatment experience.

Although there is a significant increase over time in the absolute use of EA data by payers (ρ = 0.73 and p = 0.011; **Figure 2B**) there is no evidence of a significant increase in use of EA data over time relative to the total number of appraisals conducted (ρ = 0.32 and p = 0.332).

Significant differences (χ^2= 38.8, p = 0.001) exist in the disease areas that did versus those that did not include EA data. Oncology and haematology together account for 66% of the appraisals with EA data, whereas they make up 50% of the entire fraction of appraisals. On the other hand, disease areas such as cardiology, gastroenterology, endocrinology, dermatology, rheumatology and ophthalmology jointly make up 24.5% of all appraisals, whereas they merely account for 2.6% of the appraisals that included EA data. These results can be found in **Table 1.**

Examples

To give the reader a better sense of the main labels ‘safety, efficacy, resource use’ as well as the secondary ‘treatment experience’ label, we here provide illustrative examples from the TAs that were supported by EA data.

Safety

Safety data from EA programmes are often described rather qualitatively, supporting results from clinical trials. For example, in the appraisal of gefitinib for the first-line treatment of locally advanced or metastatic non-small-cell lung cancer, the appraisal committee noted:

> *“The favourable safety profile of gefitinib demonstrated in the phase III studies is consistent with that observed in everyday settings. In addition to the data from clinical trials, the Early Access Program for*

gefitinib in Caucasian patients indicated that gefitinib is well tolerated by patients with advanced or metastatic NSCLC. The majority of ADRs associated with gefitinib are mild in nature and those most commonly reported are grade 1/2 diarrhoea and skin reactions".

Manufacturer submission, Safety and tolerability, TA192

Alternatively, safety signals from EA programmes can be quantitatively incorporated in cost-effectiveness analyses. When evaluating ocrelizumab for treating relapsing–remitting multiple sclerosis, the committee noted that an important safety signal from the compassionate use programme is lacking from the current analysis:

"The committee heard that there has been the 1 case of PML (progressive multifocal leukoencephalopathy, red.) *following treatment with ocrelizumab in the compassionate-use programme in Germany, (…). It concluded that the economic model should have included a risk of PML for ocrelizumab".*

Appraisal consultation, Adverse events in the economic model, TA533

Efficacy

Efficacy data from EA programmes can also be used, together with data from clinical trials, to estimate overall efficacy of the technology appraised. In the evaluation of lutetium (177Lu) oxodotreotide for treating irresectable or metastatic neuroendocrine tumours, response rates were obtained from the 'Erasmus study'. The Erasmus study was a compassionate use programme conducted at the Erasmus Medical Centre. The data from this programme are summarized as:

"In a single centre non-controlled phase I/II open-label study (The Erasmus study, red.)*, conducted in 810 Dutch patients with different somatostatin receptor positive tumour types, the objective response rate (ORR) for the full analysis set (FAS) population with GEP-NETs and bronchial NETs (360 patients) was 44% (95% confidence interval [CI] 38% - 49%)."*

Manufacturer submission, Executive summary, TA539

NICE requires that benefits of technologies are evaluated using quality-adjusted life years (QALYs), as NICE's decision to recommend or not recommend a product for reimbursement depends (among other things) on the willingness-to-pay for an incremental year in perfect health – the so-called cost-per-QALY approach. In the evaluation

of cabazitaxel for hormone-relapsed metastatic prostate cancer treated with docetaxel, the EA programme was used to gather quality of life data not collected during the routine clinical development:

> *"The company did not collect data on health-related quality of life in TROPIC* (the RCT, red.)*, so it took utility values from the UK Early Access Programme (EAP) for cabazitaxel. The programme measured the health-related quality of life (using the EQ-5D) of men who had been treated with cabazitaxel after docetaxel.(...)"*
>
> *"(...) One hundred and twelve patients participated in the UK EAP at 12 UK Cancer Centres. All had mCRPC with disease progression during or after docetaxel and were similar in baseline patient characteristics to the population in TROPIC. (...) Safety assessments were performed prior to each cycle and HRQL recorded at alternate cycles using the EQ-5D-3L questionnaire and visual analogue scale (VAS)."*
>
> Committee papers, Health-related quality of life, TA391

Resource use

EA data can also be used to inform other parameters in cost-effectiveness modelling. Such models are often based on Markov chains, that describe the state of the disease that patients are in at a given time point. These models require cost per state and transition probabilities or rates between states. Registries, or other real-world data sources, are frequently used to estimate such data. In the appraisal of sofosbuvir-velpatasvir-voxilaprevir for treating chronic hepatitis C, transition probabilities from decompensated liver cirrhosis to death are modelled via a Beta-distribution and the input parameters are provided from the EA programme:

> *"Variable: From decompensated cirrhosis to death*
>
> *Distribution and parameters: Beta;* $\alpha=46.5$*;* $\beta=147.2$
>
> *Source: EAP data* (expanded access programme, red.)"
>
> Manufacturer submission, Sensitivity analyses, TA507

A different, direct resource use example is given in the evaluation of ipilimumab for previously treated irresectable malignant melanoma. The dosing of ipilimumab is weight-dependent. Hence, to estimate the number of vials needed

for treatment of UK patients, an estimate of the (UK) patient population weight is required. This weight is calculated via:

> *"Patient level analysis of the weight of UK clinical trial patients in MDX010-20 (n=55), and the weight of UK patients in the ipilimumab compassionate use program (n=258), from these weights, the mean number of vials required (assuming no vial sharing) is calculated."*
>
> *"Results from these analyses showed that the dose of ipilimumab given per patient per induction has a large impact on the ICER with the minimum dose given in the trial and compassionate use programme (3 x 50 mg) resulting in an ICER of £38,387 per QALY gained and the maximum dose (2 x 200 mg) given resulting in an ICER of £88,788 per QALY gained."*
>
> Manufacturer submission, Intervention and comparators costs, TA268

Treatment experience

NHS professionals share their opinions and experience on the technology appraised in expert committee meetings. In the appraisal of patisiran for treating hereditary transthyretin amyloidosis, the Head of the National Amyloidosis Centre (NAC) is asked 'how data on real-world experience in this condition compare with clinical trial data?'. His response is:

> *"The experience of my colleagues at the NAC treating patients through compassionate access (over one year) and Early Access to Medicine Schemes has been extremely favourable. Remarkable clinically significant improvements of well-being and function have occurred in a majority of cases, including regaining the ability to walk unaided."*
>
> Clinical expert statement, HST10

Patients, caregivers, or patient group representatives are also provided the opportunity to share their experience with the appraised treatment. The assessment of nusinersen for treating muscular atrophy sparked comments from parents with children that suffer from this disease:

> *"My son is currently receiving Spinraza at Gosh for type 1c SMA. He was lucky enough to be included into the expanded access program for a select group of children. Since receiving his treatment we have watched*

the transformation of a seriously weakening child to a thriving boy who has gained significant progress in his motor function and health, we are continually amazed by his progress. He starts preschool in the coming weeks, an achievement we never thought possible. (…)"

Patient/caregiver stakeholder comment, TA588

The above provides qualitative examples of EA usage in NICE appraisals. To further illustrate how EA data are appraised by the manufacturer, ERG and NICE committee, and what the advantages and limitations of its use may be, a detailed discussion of the top-5 appraisals in which the search terms most frequently occurred can be found in the **Supplementary Files B**. This includes representative examples in the areas of haemato-oncology (e.g. prostate cancer, follicular lymphoma) and rare diseases (e.g. spinal muscular atrophy).

Discussion

In this review, we combined automated documentation searches with double, independent manual review to screen NICE documentation on the usage of EA data for HTA. We have found that data from EA programmes are frequently included: 21.1% of the TAs used EA data to evaluate safety, efficacy/effectiveness, or resource use of the appraised technology. The use of data from EA programmes appears to remain stable over the years. Additionally, patients and physicians share their treatment experience from an EA programme in 14.2% of the appraisals.

The disease areas of the appraisals that included EA data differed significantly from the overall distribution of disease areas from all appraisals investigated between 2010 and 2021. Oncology and haematology account for the lion's share (66%) of EA data usage, yet account for half (50%) of all TAs conducted. Although 'the life-threatening or seriously debilitating' prerequisite for EA is often present in haemato-oncologic malignancies, cardiac or ophthalmologic illnesses can also be severely limiting[17,18]. Cardiology and ophthalmology account for 8.4% of all TAs, but none (0%) of these programmes used EA data (or even mentioned it). There is a range of possible explanations for this discrepancy. Perhaps, drug developers in these areas may be less familiar with collecting and using EA data, or cardiologists and ophthalmologists may be less acquainted with EA than haemato-oncologists – simply because EA may be less warranted in these disease areas[19,20].

Compared with regulatory submissions to the EMA and the FDA, submissions to NICE more frequently include EA data. The EMA and FDA used EA data to support efficacy in 49 regulatory approvals over 25 years ($\mp$ 2 annually)[10].

In this work, we find that NICE used EA to inform cost-effectiveness in 76 over 11 years ($\mp$ 7 annually). One reason for this may be that payers have a higher uptake of RWD in their decision making. Furthermore, they also assess comparative effectiveness rather than efficacy. Modelling cost and comparative effectiveness by definition necessitates a variety of input parameters, every one of them potentially coming from different sources, such as EA.

Whether using EA data (or other non-randomized data) for payer decision making is wise, depends in part on the robust design and execution of the EA program, and the relevance to the decision problem[21]. The instances in which the FDA and the EMA assessed efficacy mainly based on EA data, are scarce, and characterized by (i) a high unmet medical need (ii) a rare disease population and (iii) large treatment effects[10]. Additionally, we witnessed twice (TA391, TA491) that health-related quality-of-life (HRQoL) data were not gathered during the conventional clinical trials but were captured in the EA programme. Although data from EA programme can bridge an evidence gap, HRQoL data should simply have been collected during all stages of clinical development. For safety, the use of registries, post-approval safety studies, or pharmacovigilance during EA, is useful to detect infrequently occurring adverse events. Indeed, we identified such an example in TA533, where the compassionate use programme led to the identification of a rare but serious adverse event. Overall, the evidence for assessing safety and efficacy should primarily come from regulatory studies and can be synthesized with RWD or other non-randomized sources, such as EA programmes.

Including EA can have several advantages, as it can increase sample size, add robustness, inform additional parameters - such as HRQoL - or aid to estimate effects for patients that were excluded from the trial, but were included in the EA programme. Such patients are generally older and frailer[7,22,23], and thus collecting data in these populations helps to extrapolate results on safety and efficacy found in RCTs. Estimates of resource use parameters that are derived from clinical trials, such as adherence, monitoring, or the number of hospital visits, can even be more distinct from real-world settings. Therefore, EA data can play a useful role in informing resource use parameters. Furthermore, modelling resource use requires estimates of a large number of input parameters, such as costs, incidence and also transition parameters that determine the amount of time spent in a disease state. Some of these parameters can only be estimated from studies with lengthy follow-up periods, so that patient or population registries or EA programmes would be best suited to inform decision making on these model inputs. Finally, trial values may not be sufficiently informative, as they are typically multinational and do not contain data relevant to a particular national health system.

The regulatory status of data collection during EA programmes is a matter of debate [10,11,14,15,24,25]. In Europe, individual Member States regulate EA programmes[26]. Different countries may issue conflicting statements that can be at cross with EMA decision making[10]. This also resonates in appraisals. For example, we read in the appraisal of cemiplimab for treating metastatic or locally advanced cutaneous squamous cell carcinoma:

> *"While formal data collection is not permitted from a regulatory standpoint, the safety of cemiplimab at the flat 350mg dose in a real-world setting will be monitored."*
>
> Manufacturer submission, Safety overview, TA592

This begs the questions who decides what formal and informal data collection is and whether all examples put forth in this paper where impermissible for regulators. Regardless of regulatory requirements, it can be a source of frustration when EA data are not available, as one Advisory Group (AG) noted:

> *"The lack of any efficacy data from the compassionate use program is particularly disappointing,"*
>
> AG response to company comments, AG conclusions, TA535

Although the primary intent of EA programmes is treatment provision and not to conduct research, it seems awkward to treat patients with investigational medicine and not to collect data to inform safety and efficacy. Furthermore, it is difficult to precisely determine where treatment-intent ends and research-intent starts. The changing nature of EA programmes from sole treatment-intent to treatment-intent with data collection is a current topic of debate among bioethicists[12,14,27]. We stress that data collection during EA should be light-weight and must not disproportionally burden patient and physicians – hence, a smart design should facilitate data to be collected[12]. If so, EA programmes can be the first source of RWD to inform HTA evaluations gathered in a pre-approval setting – this makes EA data different from general RWD sources (e.g. electronic health records or claims and billing data), as the latter will typically only start generating evidence once the drug has been approved. Results from EA programmes can be obtained via peer-reviewed publications, if published. Alternatively, data can be requested via the medical company using data sharing platforms, such as Vivli[28]. Finally, data may be available through local investigators (see HST7, **Supplementary Files B**).

Limitations and future research

Our work has several limitations. First, we only reviewed TAs from one HTA body: NICE. Formally, NICE's decisions are only valid within their UK jurisdiction, but informally they lead the way for other European HTA bodies - either via setting an example or via reference pricing. We have chosen NICE for our review as they have the longest history of HTA assessment and ample documentation publicly available. For other HTA bodies, results may be different. Future research should confirm whether our results uphold for other HTA bodies. Preliminary findings presented at a conference concluded that using EA data gathered within French compassionate use programmes had a positive impact on reimbursement discussions[29]. Second, we may have missed use-cases of EA data in payer submission as companies or reviewers may have used other terms to indicate EA programmes (or failed to have done so). Our automated algorithm facilitates high throughput of document screening in health policy analysis, but it may have missed cases that would have been identified in manual evaluation. Therefore, our estimates should be interpreted as a lower bound of EA use in NICE appraisals. Lastly, we were unable to exactly quantify the added value of EA data. As we lack a counterfactual, we do not know what would have happened without the inclusion of EA data. Additionally, it is difficult to measure the impact of EA data, as it is not always clear how these data have exactly been used: the use of EA data – and the appraisal thereof - in HTA by the manufacturer, ERG or NICE committee are difficult to quantify due the complexity and extent of the discussions described in the documentation. Although we have provided the reader with both high-level quantitative statistics and with illustrative qualitative examples from our data set, future research could attempt to systematically analyse these topics.

Conclusion

EA data are used in over one in five (21.1%) NICE appraisals, and this number appears to remain stable over time. In general, adding data from EA can yield more real-world information. Especially to estimate the resource use, pre-approval EA data can play a vital role informing post-approval real-world usage. In synthesis with evidence from well-controlled regulatory studies, data collected from EA programmes may meaningfully inform NICE decision making. Further research is required to understand when EA data can and should be included in health technology assessments.

Author contributions

T.B.P., D.G.J.C, J.v.R, and C.A.U. - d.G. contributed to the concept and design. T.B.P. and D.G.J.C. acquired and analysed the data. T.B.P. Drafted the manuscript. D.G.J.C, J.v.R, and C.A.U. - d.G. revised the manuscript.

Funding

This research was supported via an unrestricted grant from HealthHolland: EMCLSH20012. HealthHolland is a funding vehicle for the Dutch Ministry of Economic Affairs and Climate Policy that addresses the Dutch Life Sciences & Health sector.

Conflict of interest

C.A.U. - d.G. has received unrestricted grants from Boehringer Ingelheim, Astellas, Celgene, Sanofi, Janssen-Cilag, Bayer, Amgen, Genzyme, Merck, Glycostem Therapeutics, AstraZeneca, Roche, and Merck. T.B.P. works part-time at myTomorrows and holds stock in myTomorrows. J.v.R and D.G.J.C. declare no conflict of interest.

Data availability

The data and code from the paper are available on the github from the first author, https://github.com/TobiasPolak.

Tables

Table 1: *Technology appraisals that did ('Yes') or did not ('No') include Expanded Access (EA) data to support the profile of safety, efficacy and/or resource use, classified on disease area.*

	Included EA data			
	No[1]	Yes[1]	**Total**[1]	p-value[2]
Disease area				0.001
Benign haematology	5 (1.7%)	3 (3.8%)	8 (2.1%)	
Cardiology	14 (4.7%)	0 (0%)	14 (3.7%)	

	Included EA data			
	No[1]	Yes[1]	**Total**[1]	p-value[2]
Dermatology	12 (4.0%)	1 (1.3%)	13 (3.4%)	
Endocrinology	12 (4.0%)	0 (0%)	12 (3.2%)	
Gastroenterology	13 (4.3%)	0 (0%)	13 (3.4%)	
Haematology	34 (11%)	21 (26%)	55 (14%)	
Internal medicine	23 (7.6%)	9 (11%)	32 (8.4%)	
Neurology	14 (4.7%)	6 (7.6%)	20 (5.3%)	
Oncology	106 (35%)	32 (41%)	138 (36%)	
Ophthalmology	18 (6.0%)	0 (0%)	18 (4.7%)	
Psychiatry	3 (1.0%)	1 (1.3%)	4 (1.1%)	
Pulmonology	6 (2.0%)	4 (5.1%)	10 (2.6%)	
Rheumatology	22 (7.3%)	1 (1.3%)	23 (6.1%)	
Surgery	4 (1.3%)	1 (1.3%)	5 (1.3%)	
Urology	1 (0.3%)	1 (1.3%)	2 (0.5%)	
Vascular medicine	13 (4.3%)	0 (0%)	13 (3.4%)	
Total	300 (79%)	80 (21%)	380 (100%)	

[1] *n (%)*

[2] *Pearson chi-square test*

Table legend

Table 1: Technology appraisals that did ('Yes') or did not ('No') include Expanded Access (EA) data to support the profile of safety, efficacy and/or resource use, classified on disease area.

References

1. The National Institute for Health and Care. Guide to the methods of technology appraisal. *Nice*. 2018;(April 2013):1-93.

2. Bell H, Wailoo AJ, Hernandez M, et al. The use of real world data for the estimation of treatment effects in NICE decision making. 2016;2016(June):1-61.

3. Faria R, Hernández Alava M, Manca A, Wailoo AJ. The use of observational data to inform estimates of treatment effectiveness in technology appraisal: Methods for comparative individual patient data. *NICE Decis Support Unit*. 2015;(0):0-85.

4. National Institute for Health and Clinical Excellence (NICE). *Guide to the Single Technology Appraisal Process*.; 2009.

5. National Institute for Health and Clinical Excellence (NICE). *Guide to the Multiple Technology Appraisal Process*.; 2009.

6. Dias S, Welton NJ, Sutton AJ, Ades AE. Evidence synthesis for decision making 1: Introduction. *Med Decis Mak*. 2013;33(5):597-606. doi:10.1177/0272989X13487604

7. Blommestein H. The added value of real-world evidence. Published online 2016.

8. Sherman RE, Anderson SA, Dal Pan GJ, et al. Real-World Evidence — What Is It and What Can It Tell Us? *N Engl J Med*. 2016;375(23):2293-2297. doi:10.1056/NEJMsb1609216

9. Kimberly LL, Beuttler MM, Shen M, Caplan AL, Bateman-House A. Pre-approval Access Terminology. *Ther Innov Regul Sci*. 2017;51(4):494-500. doi:10.1177/2168479017696267

10. Polak TB, van Rosmalen J, Uyl - de Groot CA. Expanded access as a source of real-world data: An overview of FDA and EMA approvals. *Br J Clin Pharmacol*. Published online March 22, 2020. doi:10.1111/bcp.14284

11. Polak TB, van Rosmalen J, Uyl – De Groot CA. Response to Open Peer Commentary "Making It Count: Extracting Real World Data from Compassionate Use and Expanded Access Programs." *Am J Bioeth*. 2020;20(11):W4-W5. doi:10.1080/15265161.2020.1820113

12. Rozenberg O, Greenbaum D. Making It Count: Extracting Real World Data from Compassionate Use and Expanded Access Programs. *Am J Bioeth*. 2020;20(7):89-92. doi:10.1080/15265161.2020.1779857

13. Kang S, Chang S, Ross JS, Miller JE. Implementation of 21st Century Cures Act Expanded Access Policies Requirements. *Clin Pharmacol Ther*. Published online September 12, 2021. doi:10.1002/cpt.2401

14. Chapman CR, Moch KI, McFadyen A, et al. What compassionate use means for gene therapies. *Nat Biotechnol*. 2019;37(4):352-355. doi:10.1038/s41587-019-0081-7

15. Bunnik EM, Aarts N, van de Vathorst S. Little to lose and no other options: Ethical issues in efforts to facilitate expanded access to investigational drugs. *Health Policy*. 2018;122(9):977-983. doi:10.1016/j.healthpol.2018.06.005

16. Polak TB, Bunnik EM. Financial considerations in expanded access policy for gene therapies: A tough nut to crack? *Mol Ther*. Published online May 2021. doi:10.1016/j.ymthe.2021.04.030

17. Taylor DJ, Hobby AE, Binns AM, Crabb DP. How does age-related macular degeneration affect real-world visual ability and quality of life? A systematic review. *BMJ Open*. 2016;6(12):e011504. doi:10.1136/bmjopen-2016-011504

18. Chandra A, Lewis EF, Claggett BL, et al. Effects of Sacubitril/Valsartan on Physical and Social Activity Limitations in Patients With Heart Failure: A Secondary Analysis of the PARADIGM-HF Trial. *JAMA Cardiol*. 2018;3(6):498-505. doi:10.1001/jamacardio.2018.0398

19. Polak TB, Cucchi DGJ, van Rosmalen J. [Expanded Access in The Netherlands: prescribing unregistered medicine]. *Ned Tijdschr Geneeskd*. 2021;165. http://www.ncbi.nlm.nih.gov/pubmed/33651496

20. Bunnik EM, Aarts N. What do patients with unmet medical needs want? A qualitative study of patients' views and experiences with expanded access to unapproved, investigational treatments in the Netherlands. *BMC Med Ethics*. 2019;20(1). doi:10.1186/s12910-019-0420-8

21. National Institute for Health and Care Excellence. Single technology appraisal: User guide for company evidence submission. 2018;(January 2015):1-36. https://www.nice.org.uk/process/pmg24/chapter/instructions-for-companies

22. Westgeest HM, Uyl-de Groot CA, van Moorselaar RJA, et al. Differences in Trial and Real-world Populations in the Dutch Castration-resistant Prostate Cancer Registry. *Eur Urol Focus*. 2018;4(5):694-701. doi:10.1016/j.euf.2016.09.008

23. Munk NE, Knudsen JS, Pottegård A, Witte DR, Thomsen RW. Differences Between Randomized Clinical Trial Participants and Real-World Empagliflozin Users and the Changes in Their Glycated Hemoglobin Levels. *JAMA Netw Open*. 2020;3(2):e1920949. doi:10.1001/jamanetworkopen.2019.20949

24. Rawson K. Expanded Access Data Can Support Approval Decisions, US FDA Says. Accessed June 19, 2019. https://pink.pharmaintelligence.informa.com/PS124296/Expanded-Access-Data-Can-Support-Approval-Decisions-US-FDA-Says

25. S. Usdin. Beyond compassionate use: the case for using expanded access protocols to generate real world data.

26. Balasubramanian G, Morampudi S, Chhabra P, Gowda A, Zomorodi B. An overview of compassionate use programs in the european union member states. *Intractable Rare Dis Res*. 2016;5(4):244-254. doi:10.5582/irdr.2016.01054

27. Bunnik EM, Aarts N, van de Vathorst S. The changing landscape of expanded access to investigational drugs for patients with unmet medical needs: ethical implications. *J Pharm Policy Pract*. 2017;10(1):10. doi:10.1186/s40545-017-0100-3

28. Bierer BE, Li R, Barnes M, Sim I. A Global, Neutral Platform for Sharing Trial Data. *N Engl J Med*. 2016;374(25):2411-2413. doi:10.1056/NEJMp1605348

29. Villemur L, Cohignac L, Jeanbat V, Bouee S. PNS266 Collecting EARLY Access Real World Evidence DATA for a Successful Reimbursement Appraisal. *Value Heal*. 2020;23:S685. doi:10.1016/j.jval.2020.08.1710

Figure legends

Figure 1. Screening and selection of technology appraisals from NICE. STA, single technology appraisal. MTA, multiple technology appraisals. HST, highly specialized technology. EA, expanded access. NICE, National Institute for Health and Care Excellence

Figure 2: Technology appraisals (TAs) using Expanded Access (EA) data to support safety, efficacy and/or resource use. A: Venn-diagram displaying the overlap of safety, efficacy, and/or resource use labelling of TAs. **B**: Bar chart of TAs published between 01-01-2010 and 01-01-2021 that did ('Yes') or did not ('No') include data EA programmes to support safety, efficacy and/or resource use.

Supplementary Files

A. Protocol Workflow

B. Top-5 Most Referenced Appraisals

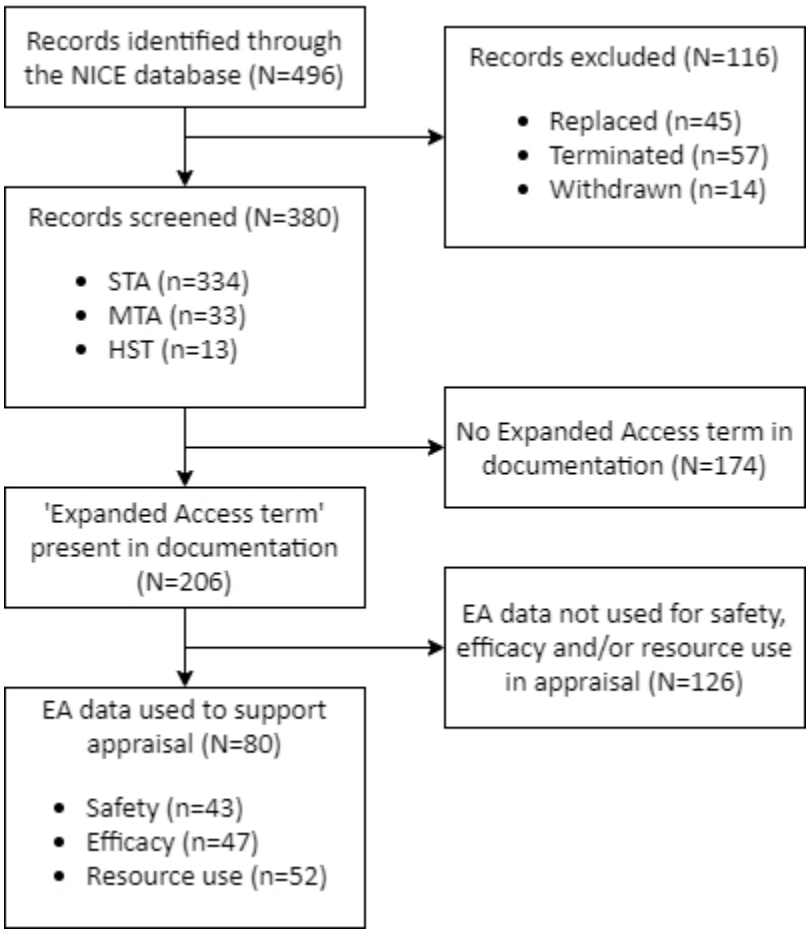


Figure 1. Screening and selection of technology appraisals from NICE. STA, single technology appraisal. MTA, multiple technology appraisals. HST, highly specialized technology. EA, expanded access. NICE, National Institute for Health and Care Excellence

141x164mm (72 x 72 DPI)

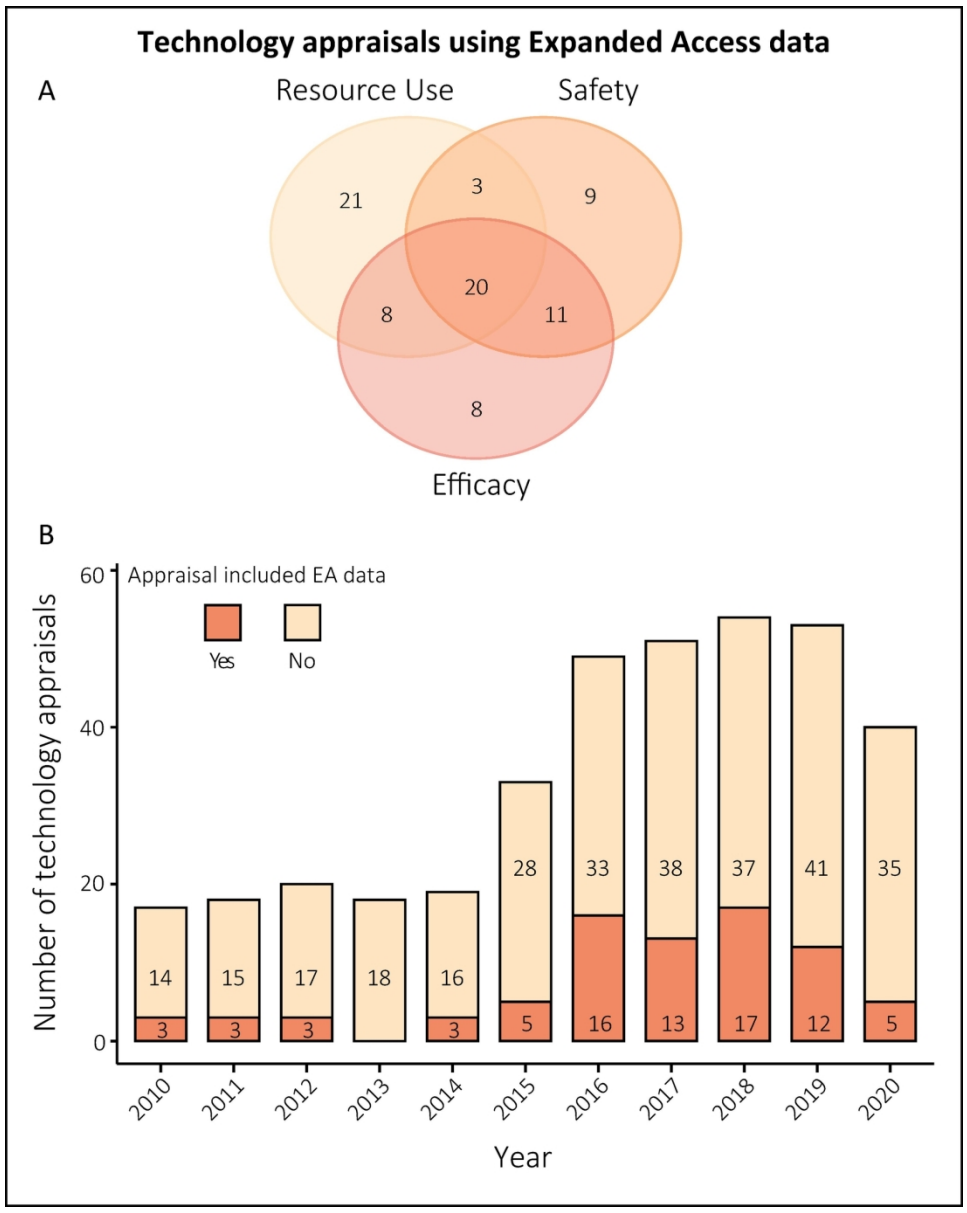


Figure 2: Technology appraisals (TAs) using Expanded Access (EA) data to support safety, efficacy and/or resource use. A: Venn-diagram displaying the overlap of safety, efficacy, and/or resource use labeling of TAs. B: Bar chart of TAs published between 01-01-2010 and 01-01-2021 that did ('Yes') or did not ('No') include data EA programs to support safety, efficacy and/or resource use.

153x191mm (300 x 300 DPI)

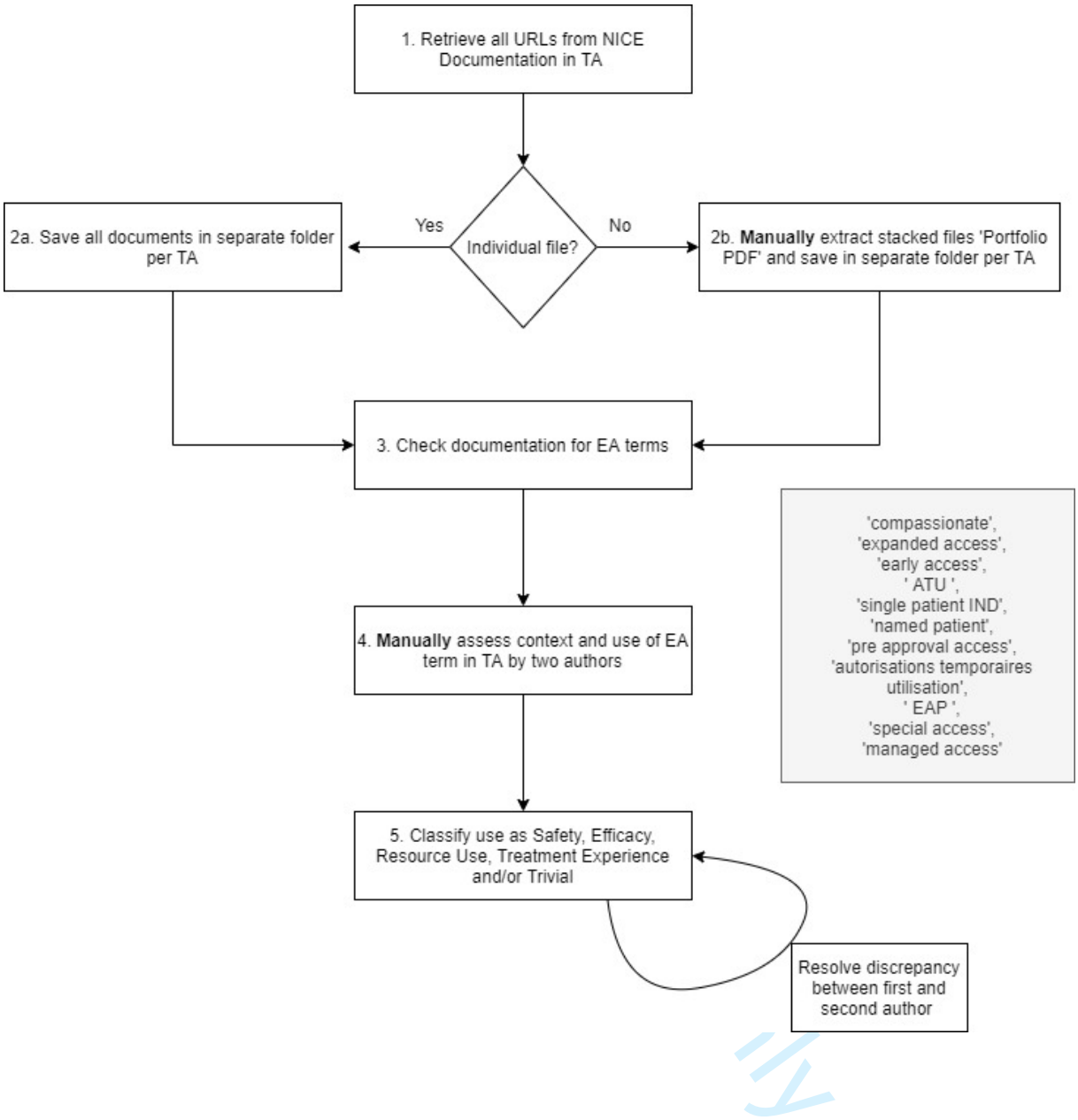
Supplementary Files to the manuscript

Real-world data from expanded access programmes in health technology assessments: a review of NICE technology appraisals

Tobias B. Polak M.Sc., David G.J. Cucchi M.D., Ph.D., Joost van Rosmalen Ph.D., and Carin A. Uyl - De Groot Ph.D.

The code and data from this paper are available at the GitHub repository of the first author: https://github.com/TobiasPolak/

A. Protocol Workflow



B. Top 5 Most Referenced Appraisals

TA391 Cabazitaxel for hormone-relapsed metastatic prostate cancer treated with docetaxel

In 2016, NICE assessed the cost-effectiveness of anticancer taxane therapy cabazitaxel for the treatment of metastatic prostate cancer that relapsed after it was treated with docetaxel. Sanofi was the submitting Company and the School of Health and Related Research (ScHARR) produced the evidence review group (ERG) report. TA391 is an updated appraisal of TA255.

In this appraisal, the company did not collect data on health-related quality of life in the main trial that investigated the use of cabazitaxel, so it took utility values from the expanded access programme in the United Kingdom. The ERG found several issues with data from this program: the open-label nature, generalizability (patients were **potentially more fit** than in the trial), the analysis was performed at interim and had not yet been subject to peer review.

The Committee partly shared the vision of the ERG: '*the Committee was concerned about the uncertainty around the utility value and whether the utility value as calculated from the early access programme could be applicable to the wider population with hormone-refractory metastatic prostate cancer refractory to docetaxel treatment'*. On the other hand, the committee also appreciated the efforts of the company: '*The committee acknowledged the limitations of using data from the UK early access programme but, in the absence of more robust evidence on health-related quality of life, it concluded that the company had used the best available data to estimate utility values'*.

The initial Final Appraisal Determination did not recommend the use of cabazitaxel, leading the company to appeal to the Appeal Panel, focusing in part on the interpretation of the EAP trial. ("the context of the EAP trial was misinterpreted, data from the EAP trial were incorrectly interpreted, and the nature of interim data was misunderstood by the committee"). The Appeal Panel '*understood both sides' positions and regarded them both as reasonable*' and as such dismissed all the grounds of appeal. After a new confidential discount to the price of cabazitaxel was arranged, its use has been recommended within the NHS.

TA667 Caplacizumab with plasma exchange and immunosuppression for treating acute acquired thrombotic thrombocytopenic purpura

In 2020, NICE assessed the cost-effectiveness of the humanized antibody caplacizumab, used together with plasma exchange and immunosuppression for the treatment of acute acquired thrombotic thrombocytopenic (TTP) purpura. Caplacizumab inhibits the interaction between Von-Willebrand-factor and thrombocytes, thereby reducing the aggregation of thrombocytes which is typical for TTP. Sanofi was the submitting Company and the Peninsula Technology Assessment Group (PenTAG) produced the evidence review group (ERG) report.

In the main trial for caplacizumab (HERCULES, N=145), no patient died while on treatment with caplacizumab (0%). Due to the unreliability of mortality data from the trial (clinicians noted that mortality was unlikely to be 0%), data from the compassionate use programme was brought in. At the first data lock, 8 out of 187 (4.2%) of the patients perished, and 9/239 (3.8%) at second data lock.

The limited information available rendered the interpretation of these data difficult. Mortality from the compassionate use programme was based on deaths reported via Adverse Event Reporting. No baseline characteristics were available to compare patients among data sources: *The monitoring programme for caplacizumab was a compassionate use programme rather than a data collection programme. As such, the only information available includes where the patient was from, whether caplacizumab was received and whether the patient died, (…). Therefore, an assessment of the similarity between mortality sources using patient characteristics could not be conducted .*

Therefore, the ERG '*notes potential ambiguities and sources of bias in the compassionate use program (…) including unknown follow-up periods, unclear recruitment process, and that it draws from an international population'*.

The company interjected that '*the compassionate use programme estimates selected to represent caplacizumab in the comparison are, if anything, too high*' – as *'clinicians agreed that treatment with caplacizumab is started later in the compassionate use programme that it would be if it was made available through routine funding (as requests are individual and caplacizumab is not available on site). Mortality data based on this programme should therefore be considered as the maximum mortality expected with caplacizumab.'*

The committee agreed that it was impossible to '*estimate reliably the extent of the benefit using the randomised trial data*' and recognized the need for use of data on deaths from the global compassionate use scheme. It noted that the absolute rate of death for people treated with caplacizumab under the compassionate use scheme was likely to be valid, but that the relative benefit ascribed to caplacizumab from observational data '*was very likely to be confounded'*.

Furthermore, the committee noted that *'Some potential cost savings associated with caplacizumab may not be included in the company's model'* as '*The company stated that, based on its observations from the compassionate use scheme for caplacizumab, in NHS clinical practice, people would have it for a shorter duration than in the trials. The committee in general prefers not to disassociate estimates of cost and effectiveness from a trial. However, it appreciated that many assumptions about caplacizumab's effectiveness in this model were not taken from the main trial. It also thought that some potential cost savings associated with caplacizumab may not have been included in the company's model.'*

Despite the remaining uncertainty, *'(…) the assumptions in the economic modelling are plausible. Also, there are potential benefits with caplacizumab that are not included in the cost-effectiveness estimates. Overall, the estimates are within the range normally considered a costeffective use of NHS resources. So, caplacizumab is recommended for treating acute acquired TTP'.*

TA588 Nusinersen for treating spinal muscular atrophy

In 2018-2019, NICE assessed the cost-effectiveness of the antisense oligonucleotide nusinersen, used in the treatment of spinal muscular atrophy. Nusinersen promotes the formation of the functional SMN protein, through modulation of intron splicing, essential for normal function of motor neurons. Sanofi was the submitting Company and the Peninsula Technology Assessment Group (PenTAG) produced the evidence review group (ERG) report.

NICE initially did not recommend the use of nusinersen for treating SMA as it was not deemed a cost-effective use of NHS resources. NICE consulted with the public and professionals and noted that '*Following consultation, the committee heard that there was real-world evidence that would be relevant for the committee's decision making that had not been considered by the company.'*

Although the Company briefly touches upon data from the early access programme (EAP) in UK and Ireland (63 patients, of which 25 males and 38 females) and additionally points at the publication of a second European EAP conducted in other European countries (N=36, Gargaun et al.), the Spinal Muscular Atrophy Support UK and The SMA Trust points to several other studies in the consultation period: '*We note that the real-world studies only review outcomes for children with SMA Type 1 for the first six months of treatment but consider 'real world' evidence critical to decision making. They all assist with confirming the certainty of evidence of effectiveness (see below). In particular we refer to: Reviews of the Expanded Access Programme:*

- *Europe - 33 children aged from 8.3 to 113.1 months - December 2016 - May 2017. Aragon-Gawinska, K et al. (2018)*
- *Australia – 16 patients aged 2.5 months to 35.7 years November 2016 – September 2017 Farrar, M et al. (2018)*
- *England - Great Ormond Street Hospital – 21 patients aged 8.3 – 113.1 months March – October 2017 Tillmann, A et al. (2018)*
- *Germany – 61 patients aged 1 – 93 months in seven neuromuscular centres November 2016 – June 2017 Pechmann, A et al. (2018)*
- *Italy – 104 patients – aged 3 months – 19 years 9 months - first six months of EAP Pane, Pane M et al. (2018)*
- *Hoy, S (2018)'*

The committee responded: '*The committee have taken into account the consultation comments including the views of patients, carers and clinical experts alongside the updated economic model and proposed MAA. Nusinersen is now recommended for pre-symptomatic and types 1, 2, and 3 SMA in the context of a MAA*.

The company stated that they did not consider these data '*because the results were consistent with the clinical data that it had presented and, in comparison, the data were immature, would be from non-UK sources and would only include SMA type 1'.* The committee stated '*that it would have liked the company to identify supportive real world evidence, given the clinical uncertainties identified.'* - but also acknowledged that the company already included several types of data.

In the end, Nusinersen became available through a managed access agreement *', including the collection of more data to address the uncertainties. '*

HST7 Strimvelis for treating adenosine deaminase deficiency–severe combined immunodeficiency

In 2018, NICE assessed the cost-effectiveness of strimvelis, used to treat severe combined immunodeficiency due to adenosine deaminase deficiency (ADA-SCID). Patients with ADA-SCID have a dysfunctional gene, needed for the production of the enzyme adenosine deaminase (ADA), leading to defective lymphocytes and thereby severe immunodeficiency. Strimvelis consists of genetically modified bone marrow cells of the patient, reactivating ADA production. Strimvelis is used in patients who are ineligible for allogeneic bone-marrow transplantation. Since strimvelis is a gene therapy product, its cost-effectiveness is evaluated through a 'highly specialised technology guidance' (HST). GlaxoSmithKline was the submitting company, and the Centre for Reviews and Dissemination and Centre for Health Economics in York prepared the Evidence Review Group (ERG) report.

The submission of data for the gene therapy Strimvelis comprised a mix of evidence sources: '*The safety and efficacy of Strimvelis have been evaluated in a programme comprising 2 pilot studies, 1 pivotal study, a compassionate use programme (CUP), and a long-term follow-up (LTFU) study.'*

The company preferred to report the results of the clinical trials together, as an '*integrated population', with results from the Named Patient Programme (NPP) presented alongside as supportive evidence. The company stated that it did not include the NPP data in the integrated population because the population of the NPP was substantially different to the population in the other trials, and that it could not access all the patient-level data because the NPP was a clinician-initiated process.*

The ERG critiqued this decision: '*However, the ERG did not consider it appropriate that data from the Named Patient Population were excluded from the narrative synthesis of clinical effectiveness evidence. This is particularly important given the small sample size of the Strimvelis Integrated Population (n=18) and therefore the need to consider all available data when evaluating the effectiveness of this treatment.'*

Indeed, the Company even requested (to no avail) the ERG to remove the wording 'NPP study', as '*Noting the NPP as a study wrongly indicates that the NPP is part of the Strimvelis clinical programme and therefore at the same level in terms of availability and quality of evidence.*

NICE however specifically requested more information on these patients. '*A3. Please provide a narrative summary of the data (e.g. in terms of overall survival, intervention-free survival, adverse events etc.) available from the named patient programme using the same format as in the main clinical effectiveness section on the Strimvelis Integrated Population.'*

As the named-patient program was investigator-initiated, access to data was limited: 'A3. *Table 1 contains the requested information, as available, for patients in the NPP. Data on the proportion of patients with viral infection at baseline are not available. As the ERG has noted, the NPP is not run by GSK, which limits access to data and as such it is difficult to speculate on wider applicability of these immature and incomplete data. The programme is ongoing and data are not scheduled for formal analysis until all patients have reached 3 years of follow-up'.*

Strimvelis is recommended as a treatment option for treating adenosine deaminase deficiency–severe combined immunodeficiency (ADA–SCID) when no suitable human leukocyte antigen-matched related stem cell donor is available.

TA604 Idelalisib for treating refractory follicular lymphoma

In 2019, NICE assessed cost-effectiveness of idelalisib, used as monotherapy for refractory follicular lymphoma, a malignancy of B-lymphocytes. Idelalisib is a kinase inhibitor, reducing the activity of phosphoinositide 3-kinase p110δ (PI3Kδ), which is an enzyme involved in growth, proliferation, differentiation and survival of blood cells. PI3Kδ is known to be overactive in B-cell malignancies, and is therefore used as therapeutic target in follicular lymphoma. Gilead was the submitting company, and Kleijnen Systematic Reviews produced the Evidence Review Group (ERG) report.

The single-arm main trial (DELTA) was supplemented with data from the Compassionate Use Program: *The company supplemented the DELTA study with another source of evidence for idelalisib: the Compassionate Use Programme (CUP). This provided retrospective observational data from patients with follicular lymphoma having compassionate treatment in the UK and Ireland. The company took a subset of 79 patients with relapsed or refractory follicular lymphoma that had been treated with idelalisib. In these patients, median progression-free survival was 7.1 months, and median overall survival was not reached.*

The Committee decided that **neither** data set was '*adequate enough for using to determine how well patients on idelalisib fared compared with people who had not taken idelalisib.'* Despite the absence of controlled trials, the committee discussed the evidence presented to determine which set (trial or CUP) was most generalizable to the use of idelalisib clinical practice. Evidence was ambivalent:

- *The committee noted the difference in Eastern Cooperative Oncology Group (ECOG) performance status and Follicular Lymphoma International Prognostic Index (FLIPI) I and II scores between DELTA and the CUP. Notably, 8% of patients in DELTA had an ECOG score of 2 to 4 compared with 25% of patients in the CUP, reflecting poorer performance among patients in the CUP. The clinical experts stated that the ECOG performance status in CUP more closely reflected clinical practice than that in DELTA.*
- *The clinical experts noted that the time since completing the last therapy was shorter in DELTA than in the CUP, suggesting that patients in DELTA had a poorer prognosis*

Resulting in the ambivalent conclusion that *'the populations in DELTA and the CUP were different. (…) Also, patient and disease characteristics at baseline differed, with some suggesting a more favourable prognosis in DELTA than in the CUP, and others suggesting the opposite.' Even help from clinical experts could not resolve the issue, as 'the clinical experts suggested that the CUP cohort was more likely to reflect the intended UK treatment population because it was a 'real-world' study with patients from Britain and Ireland. However, the clinical experts acknowledged that such studies lack the methodological rigour typical of a clinical trial.'* In the end, '*The committee concluded that it was unclear whether the DELTA population or the CUP cohort more closely reflected clinical practice and took both into account for decision making.'*

Idelalisib was not recommended in the Final Appraisal Determination. '*There are a wide range of cost-effectiveness estimates but, because the evidence is weak, idelalisib is not considered to represent a cost-effective use of NHS resources. Therefore, idelalisib cannot be recommended for routine use in the NHS.'*

STROBE 2007 (v4) Statement—Checklist of items that should be included in reports of *cross-sectional studies*

Section/Topic	Item #	Recommendation	Reported on page #
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract	1
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	2
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	3
Objectives	3	State specific objectives, including any prespecified hypotheses	3
Methods			
Study design	4	Present key elements of study design early in the paper	4
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	4
Participants	6	(*a*) Give the eligibility criteria, and the sources and methods of selection of participants	4
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	4
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	4
Bias	9	Describe any efforts to address potential sources of bias	8
Study size	10	Explain how the study size was arrived at	4
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	4
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding	4
		(*b*) Describe any methods used to examine subgroups and interactions	4
		(*c*) Explain how missing data were addressed	N/A
		(*d*) If applicable, describe analytical methods taking account of sampling strategy	N/A
		(*e*) Describe any sensitivity analyses	N/A
Results			

Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	4
		(b) Give reasons for non-participation at each stage	4
		(c) Consider use of a flow diagram	4
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	5
		(b) Indicate number of participants with missing data for each variable of interest	5
Outcome data	15*	Report numbers of outcome events or summary measures	5
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	5
		(*b*) Report category boundaries when continuous variables were categorized	6
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	N/A
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	N/A
Discussion			
Key results	18	Summarise key results with reference to study objectives	6
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	8
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	7
Generalisability	21	Discuss the generalisability (external validity) of the study results	8
Other information			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	10

*Give information separately for cases and controls in case-control studies and, if applicable, for exposed and unexposed groups in cohort and cross-sectional studies.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at www.strobe-statement.org.