



Fixed effect versus random effects Bayesian network meta-analyses in practice: what is used to inform National Institute for Health and Care Excellence (NICE) technology appraisals?

Cheah Z¹, Gittfried AMV², Ainsworth C¹

¹OPEN Health HEOR & Market Access, London, UK
²OPEN Health HEOR & Market Access, Rotterdam, The Netherlands

BACKGROUND AND OBJECTIVES

- Literature, including NICE Technical Support Document (TSD) 2¹, provides clear guidance on the use of fixed effect (FE) and random effects (RE) assumptions in pairwise and network meta-analyses (NMA).

FE is preferred	RE is preferred
When it is appropriate to believe that the true treatment effect estimated by studies is the same	When there is evidence (either statistical or clinical) of heterogeneity in treatment effects between studies

- However, implementation in practice may differ, with continued use of FE models even when heterogeneity is present within the evidence base.²
- The aim of this study was to review NICE technology appraisals (TAs) to assess the application of FE and RE assumptions in Bayesian NMAs supporting company submissions (CSs).

METHODS

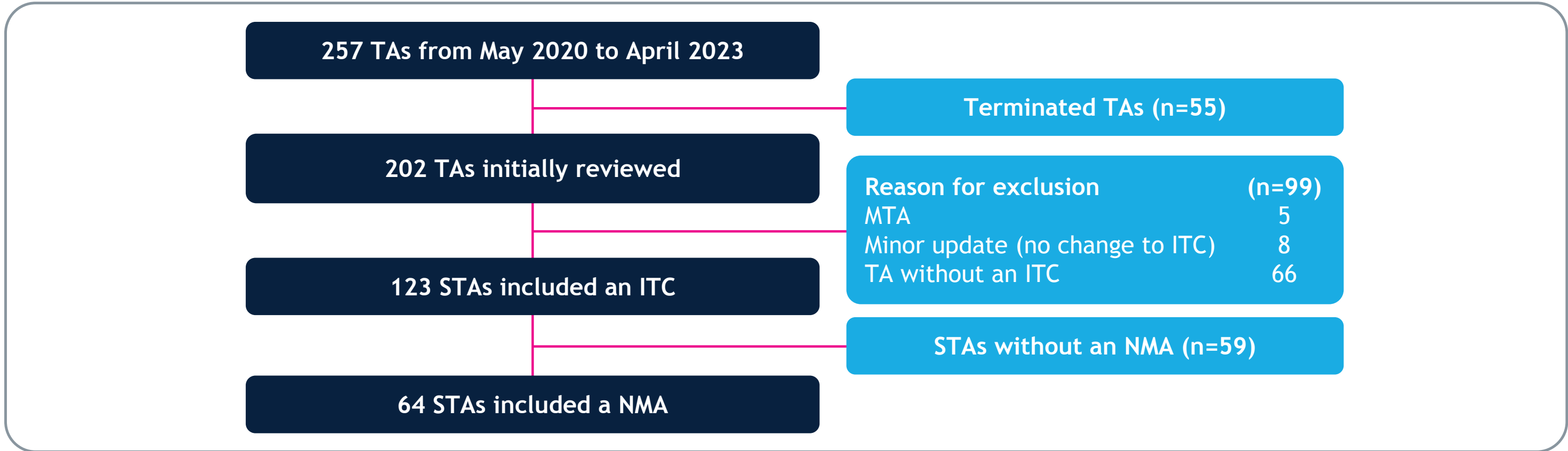
- TAs published on the NICE website between May 2020 and April 2023 were considered.
 - Multiple technology appraisals (MTAs), TAs with changes to marketing authorisation, recommendation updates, or revised patient access schemes (rather than resubmissions) and TAs without a supporting indirect treatment comparison (ITC) were all excluded.
- Where reported in publicly available committee papers, the following information was extracted: use of FE and RE models, company and external assessment group (EAG)/evidence review group (ERG) model selection, and prior distribution used for the between-study standard deviation (SD) (where RE models were reported).
- Requests for unavailable supporting documents were not made.

RESULTS

TA identification process

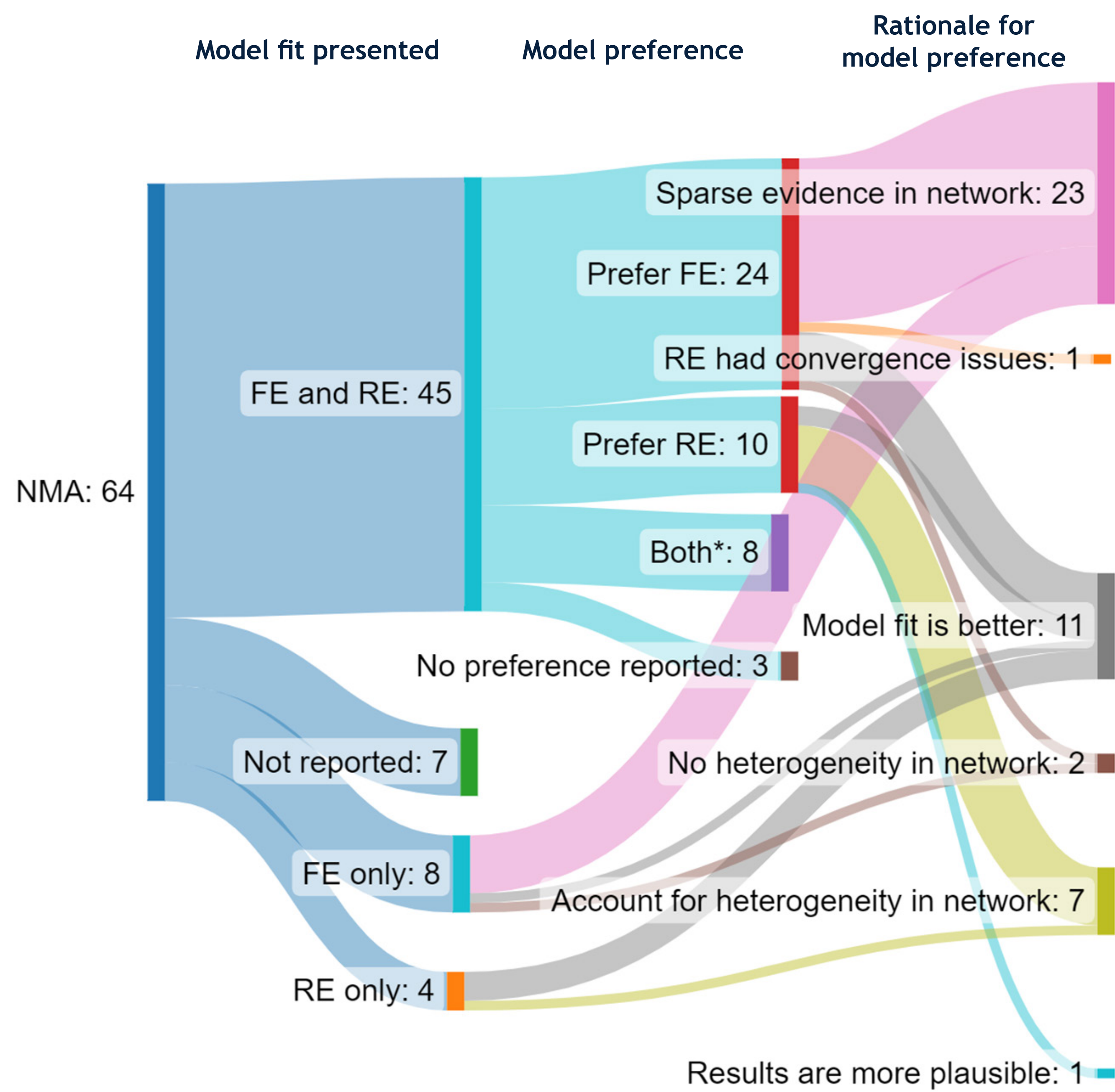
- Of the 257 TAs identified, 64 included a Bayesian NMA in the CS and were included (Figure 1).

Figure 1: Flow chart showing the identification of NICE TAs including an NMA



Abbreviations: ITC, indirect treatment comparison; MTA, multiple technology appraisal; NMA, network meta-analysis; STA, single technology appraisal; TA, technology appraisal.

Figure 2: Sankey diagram outlining CS model selection preference and justification



CS implementation of FE or RE models

- Of the 64 TAs that included an NMA, the majority (70.3%) of CSs fitted both FE and RE models (Figure 2), with few (18.8%) performing NMAs using only one assumption for treatment effects (12.5% and 6.3% for FE and RE, respectively).
- For the remaining TAs (10.9%), the type of model fit in the CS was not reported, nor was it clear based on reading additional materials available in the published TA.

CS model selection preference and justification

- For the 45 CSs that fit both FE and RE models, 53.3% and 22.2% preferred presentation of NMA results under a FE and RE model, respectively (Figure 2).
- Further, of all 32 CSs that performed analyses using only a FE model or preferred the presentation of results under a FE model:
 - Most (71.9%) reasoned the choice due to sparse evidence to inform a network, and subsequent lack of data to inform estimation of a RE model or between-study heterogeneity.
- Of the 14 CSs that performed analyses using only a RE model or preferred a RE model:
 - Most (50.0%) reasoned the choice as accounting for heterogeneity being present.
 - Whilst 42.9% reasoned similar or better fit compared to the FE model.

EAG/ERG model selection critique and preference

- Table 1 illustrates the disparity in model selection preference between the manufacturer and EAG/ERG for the 45 CSs which fit both FE and RE models.
- Of the 24 CSs where the manufacturer preferred presentation of NMA results under an FE model, the EAG/ERG only agreed in 16.7% of cases.
- Conversely, the EAG/ERG agreed with 40.0% of CSs that preferred a RE model.
- Further, in 16.7% and 20.0% of CSs that preferred a FE model and RE model, respectively, the EAG/ERG preferred or recommended RE models using an informative prior for the between-study SD.
- EAG/ERG preference was not reported or unclear in a substantial number (44.2%) of TAs.

Table 1: Manufacturer versus EAG/ERG model selection preference

	EAG preference						
		Both*	FE	RE	NR	RE with informative prior	Unclear
Manufacturer preference	Both*	3	0	3	1	0	0
	FE	1	4	3	3	4	9
	RE	0	1	4	2	2	1
	NR	0	0	0	4	0	0

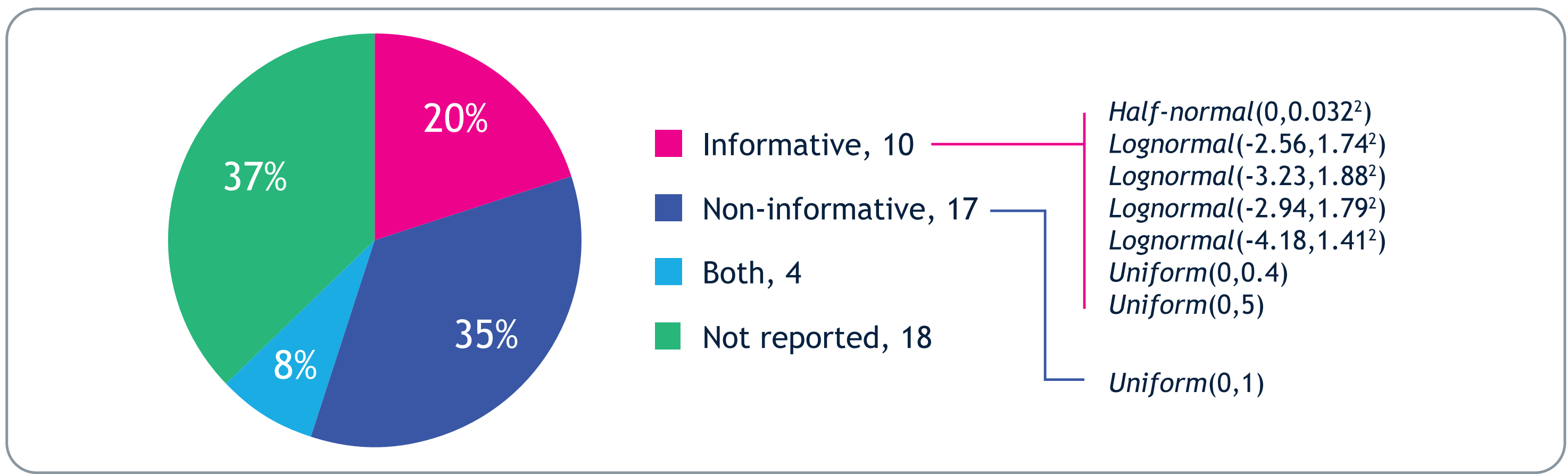
Note: Values represent number of TAs.

*Both FE and RE were fit for different outcomes by the company and the EAG/ERG agreed with the model fits.

Priors for between-study SD

- Of the 49 TAs fitting a RE model, 63.3% included details of the prior distribution used for the between-study SD in the CS (Figure 3).
 - Approximately 20.0% used informative priors, frequently using a lognormal distribution in line with guidance from Turner et al., (2012)³.
 - Non-informative (or vague) priors were used in 35.0% of submissions.
 - Most (94.1%) CSs using non-informative priors did not report the prior distribution used. Where reported, the implemented prior followed a uniform distribution.

Figure 3: Summary of between-study SD prior distributions used in CSs fitting RE models



LIMITATIONS

- The primary limitations of this study were the 3-year time horizon and the reliance on publicly available information from NICE TAs.
- It is unclear whether lack of details on FE and RE NMA is due to redaction or unpublished appendices, or because such information was not provided to begin with.

CONCLUSIONS

- EAG/ERG preference for model selection was not reported or unclear in a substantial number of TAs, making it difficult to determine manufacturer and EAG/ERG alignment.
 - Even so, differences were observed between manufacturer and EAG/ERG preference.
- Results of this study were in line with previous research², finding that FE is being overused, particularly in small networks, using the justification that there is insufficient evidence to use RE, while the primary reason for FE use should be a lack of heterogeneity.
 - This suggests that companies may not be appropriately accounting for heterogeneity.
- Informative priors were also not frequently used, although recommended by guidelines and EAG/ERG, indicating potential for improved convergence in RE models and increased use.
- This study concludes that FE overuse has been and continues to be a problem in TAs, and as the issue persists it should be clearly restated and enforced by NICE. Furthermore, clarity in EAG/ERG feedback on model selection could be improved.

REFERENCES

- Dias et al. NICE DSU TSD; 2. NICE; 2014; 2. Ren et al. *Med Decis Making*. 2018;38(4):531-542; 3. Turner et al. *Int J Epidemiol*. 2012;41(3):818-827.

ACKNOWLEDGEMENTS

We want to thank Maarten Treur for his review of this poster, and Ana Moura for her contributions conceptualizing this project.