Review of evidence underpinning recommendations in past National Institute for Health and Care Excellence HealthTech guidance

Carr E,[1] Macmillan T,[1] Walker T [2]

1 York Health Economics Consortium, University of York, York, YO10 5NQ

2 National Institute of Health and Care Excellence

Poster number
HTA133

INTRODUCTION

The National Institute for Health and Care Excellence's (NICE) health technology guidance (HTG) programme evaluates health technologies to determine clinical and cost effectiveness for the NHS. Based on the available evidence and analysis, technologies are given **one of four possible recommendations**:

1. **Can be used**: Recommends routine adoption in the NHS.
2. **Can be used during evidence generation**: Conditional use while further data is gathered.
3. **More research needed**: Restricts use strictly to formal research settings.
4. **Should not be used.**

The objective of this review was to evaluate the evidence used to support recommendations in the HTG programme and how the availability and quality of evidence influences recommendation decisions.

METHODS

We selected HTG documents and reviewed the evidence base used to support different recommendations. We focused on the **medical technologies evaluation programme (MTEP) and the diagnostics assessment programme (DAP).**

Documents published between 2019 and 2026 were identified using their old URLs (e.g. https://www.nice.org.uk/guidance/dg50), in descending chronological order of the original NICE MTG or DG number, until a sample of 50 guidelines had been gathered.

For each guidance document we extracted the following characteristics and assessed any patterns or connections with the recommendations made:

- Study design and sample size
- Results of, and evidence informing, economic models
- Study applicability and generalisability
- Clinical, safety and cost outcomes
- Quality and risk of bias assessment

CONCLUSIONS

As expected, a finding of cost-effectiveness was the characteristic that most consistently aligned with a positive recommendation across the 50 guidance documents. Larger study sample sizes appeared to be associated with positive recommendations. For DAP guidelines, concerns about the applicability of the evidence base to the decision problem were more common for diagnostics that received research only or no recommendation.

RESULTS

50 guidelines were included making a total of 63 recommendations across devices and indications (**Figure 1**).

- **MTEP (34 guidelines, 37 recommendations)**: 18 fully recommended, 4 can be used in evidence generation period, 14 research recommendation only, 1 not recommended.
- **DAP (16 guidelines, 26 recommendations)**: 2 can be used in evidence generation period, 9 fully recommended, 13 research recommendation only, 2 not recommended.

No relationship was identified between risk of bias or safety concerns in the evidence base (in the latter case generally due to no safety concerns being identified in most reviews) and final recommendation. The other evidence base characteristics are discussed below.

Study design

Among MTEP guidelines there were no clear connections between the highest level of evidence in terms of study design and the type of recommendation. However, for the 26 DAP guidelines there was a tendency for diagnostics supported by randomised controlled trials (RCTs) to receive positive recommendations:

Table 1: Availability of RCT evidence across DAP recommendations

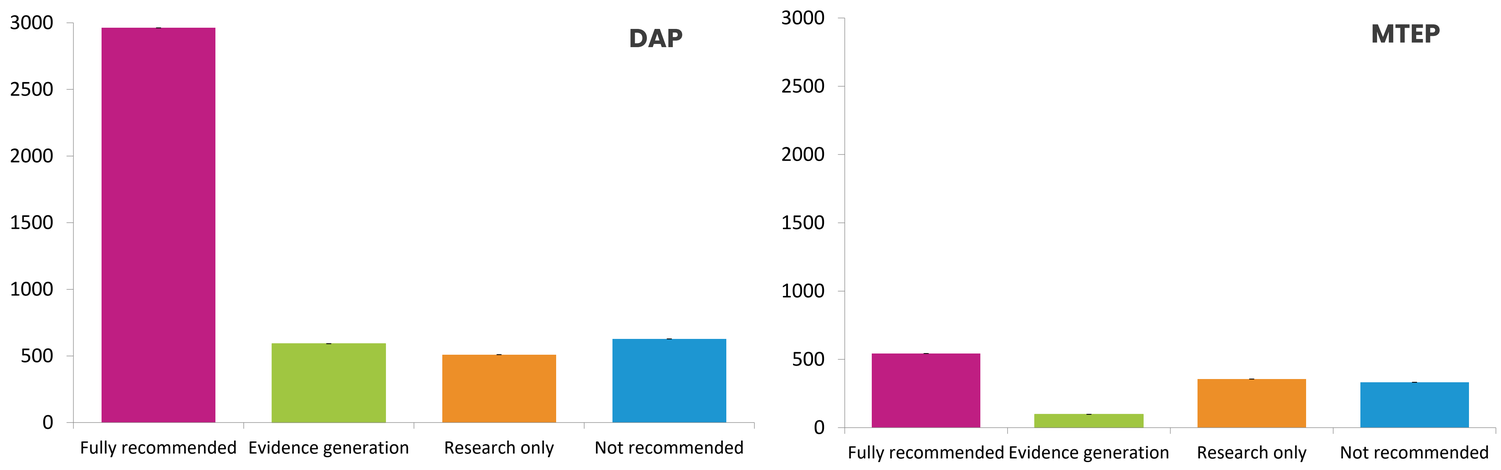
DAP recommendations	RCT evidence available	No RCT evidence available
Full (n=9)	77%	23%
Evidence generation (n=2)	50%	50%
Research only or not recommended (n=15)	33%	67%

DAP: diagnostic assessment programme; RCT: randomised controlled trial

Sample size

In both the MTEP and DAP guidelines, devices or diagnostics that received full recommendations tended to be supported by larger studies (**Figure 2**). However, technologies that received other types of recommendation were not characterised by especially small sample sizes; median sample sizes were >300 patients for all recommendation types across MTEP and DAP other than MTEP evidence generation recommendations (median 97.5 patients).

Figure 1: Median study sample size, MTEP and DAP



Applicability and generalisability

Applicability and generalisability concerns were not associated with recommendation type in the MTEP guidelines. There was a greater tendency for the committee or EAG to raise concerns in DAP guidelines that received research only or no recommendation:

Table 2: Applicability/generalizability concerns across DAP recommendations

DAP recommendations	No concerns	Concerns
Full (n=9)	77%	23%
Evidence generation (n=2)	0%	100%
Research only or not recommended (n=15)	40%	60%

DAP: diagnostic assessment programme

Cost effectiveness

A finding of cost effectiveness was the characteristic that most consistently aligned with a positive recommendation. All full or evidence-generation recommendations in both MTEP and DAP guidance reviews were made following a finding of cost-effectiveness with only three exceptions (HTG651, HTG657, HTG708). Only one technology was not recommended despite a finding of cost-effectiveness (HTG649).

York Health Economics Consortium

E: emma.carr@york.ac.uk
W: www.yhec.co.uk
T: +44(0) 1904 323 620