BMJ OPEN

Externally peer reviewed? Yes
Evidence type: Data analysis
Subjects: Health Technology Appraisals

Poor quality clinical data informing NICE decisions on treatments in over half of cases

20 year analysis shows no improvement in data quality submitted by manufacturers

The quality of evidence submitted to the National Institute for Health and Care Excellence (NICE) for informing its decisions to recommend technologies for use in the NHS was poor in more than half of cases, reveals a 20-year analysis, published in the open access journal ***BMJ Open.***

And the data quality submitted for health technology appraisals by manufacturers between 2000 and 2019 was consistently poor, with no improvement during that time, the analysis shows.

NICE advises the NHS on the clinical and cost-effectiveness of both new and existing technologies, through an independent body of specialists, which make up its appraisal committee.

This committee's decisions are based on reports from its independent technology assessors, plus advice from consultees, clinical, NHS commissioning, and patient experts.

In recent years, NICE has expanded the evidence it considers, to include a broader range of factors that influence health: registry data; national statistics; surveys; clinical practice recommendations; expert opinions; and additional knowledge from manufacturers.

In light of this, the researchers wanted to systematically review all NICE's active technology appraisals published between 2000 and 2019 to scrutinise the clinical evidence submitted by the manufacturers and assess its quality for decision-making.

They therefore extracted data from the independent assessment group and evidence review group reports and final appraisal determinations on the quality of submitted randomised controlled clinical trials and the overall quality of evidence submitted for decision-making.

For single technology appraisals (STAs), which evaluate a single product, device, or technology for a single indication, and usually involve new drugs or indications, they also extracted data on quality of life evidence and comparative clinical evidence.

Each category was scored for quality—2 for good; 1 for acceptable; 0 for poor; and −1 for unacceptable. The scores were peer reviewed by all the members of the research team to try and minimise bias.

In all, the evidence for 409 technology appraisals was analysed: 104 multiple technology appraisals (MTA), which evaluate technologies that share one or more criteria; and 305 STAs.

The appraisals included 25 non-pharmaceutical products, 14 medical devices, 6 other therapies, 5 surgical procedures, and 384 drugs.

In two thirds of all appraisals, the overall quality of evidence was judged to be either poor (224; 55%) or unacceptable (41;10%). The quality of evidence was judged acceptable in a third (139, 34%) or good in only 1% (5).

In nearly 4 out of 10 (39%; 119) STAs, the quality of comparative evidence was considered poor, and in 17% (51) unacceptable. In 44% (135) the quality of quality of life data was considered poor, and unacceptable in 15% (47). In only a third (102) of STA appraisals was the quality deemed acceptable, and good in only 7% (21).

Based on analysis of the comments from the review group reports, over half of the clinical trials presented in the manufacturers' submissions for all appraisals were deemed to be either poor (166; 41%) or unacceptable (40;10%) quality. Just under half were considered to be of acceptable (173; 42%) or good (30;7%) quality.

Weak or insufficient evidence from poorly conducted clinical trials was often used because it was the only evidence available. But even when the trials had been done well, and the evidence was comprehensive, the comparators were often unsuitable for decision-making in the NHS context, the analysis showed.

And over the entire 20 year period the overall quality of evidence submitted to NICE didn't change, and was consistently poor, the analysis indicated.

In particular, the researchers noted a lack of clarity on the methodologies used by the manufacturers when carrying out systematic reviews and indirect comparisons.

They found that comparator data often didn't reflect the UK population and routine treatment pathways. Indirect comparisons were used in over two thirds (68%;207) of STAs to establish the comparative clinical effectiveness of interventions.

And the quality of life data was often of poor or unacceptable quality, even if collected in pivotal trials; and clarity in reporting methodology and details by both manufacturers and assessment bodies varied significantly.

The researchers acknowledge various limitations to their findings, chief among which was the subjective scoring system used for grading the quality of evidence, and the focus on only certain elements of the evidence submitted by manufacturers.

But they nevertheless conclude: "We found that the primary components of clinical evidence (comparative clinical effectiveness, measures of [quality of life] outcomes and overall design of [randomised controlled trials]) that influence patients and are crucial for NICE's decision making framework are of poor quality.

"Since the evidence bar continues to be lowered, it is essential to have [health technology assessment] bodies and payers' input to ensure that the generation of evidence submitted to NICE is strengthened.

"However, it is essential that stakeholders are aware of this and that organisations put more effort into generating high-quality evidence premarket and postmarket entry."