VALUE IN HEALTH ▮ (2017) ▮▮▮–▮▮▮

Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/jval

The Role of Noncomparative Evidence in Health Technology Assessment Decisions

Elizabeth A. Griffiths, BSc[1,], Richard Macaulay, BSc, PhD[1], Nirma K. Vadlamudi, BA, BS, MPH[2], Jasim Uddin, BSc, PhD[1], Ebony R. Samuels, BSc, PhD[1]*

[1]*PAREXEL Access Consulting, PAREXEL International, London, UK;* [2]*Faculty of Pharmaceutical Sciences, University of British Columbia, Vancouver, British Columbia, Canada*

ABSTRACT

Background: Many health technology assessment (HTA) agencies express a preference for randomized controlled trial evidence when appraising health technologies; nevertheless, it is not always feasible or ethical to conduct such comparative trials. **Objectives:** To assess the role of noncomparative evidence in HTA decision making. **Methods:** The Web sites of the National Institute for Health and Care Excellence (NICE) in the United Kingdom, the Canadian Agency for Drugs and Technologies in Health (CADTH) in Canada, and the Institute for Quality and Efficiency in Health Care (*Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen* [IQWiG]) in Germany were searched for single HTA reports (published between January 2010 and December 2015). The product, indication, outcome, and clinical evidence presented (comparative/noncomparative) were double-extracted, with any discrepancies reconciled. A noncomparative study was defined as any study not presenting results against another treatment (including placebo or best supportive care), regardless of phase or setting, including dose-ranging studies. **Results:** A total of 549 appraisals were extracted. Noncomparative evidence was considered in 38% (45 of 118) of NICE submissions, 13% (34 of 262) of CADTH submissions, and 12% (20 of 169) of IQWiG submissions. Evidence submissions based exclusively on noncomparative evidence were presented in only 4% (5 of 118) of NICE appraisals, 6% (16 of 262) of CADTH appraisals, and 4% (6 of 169) of IQWiG appraisals. Most drugs appraised solely on the basis of noncomparative evidence were indicated for cancer or hepatitis C. Positive outcome rates (encompassing recommended/restricted/added-benefit decisions) for submissions presenting only noncomparative evidence were similar to overall recommendation rates for CADTH (69% vs. 68%, respectively), but were numerically lower for NICE (60% vs. 84%, respectively) and IQWiG (17% vs. 38%, respectively) ($P > 0.05$ for all). **Conclusions:** Noncomparative studies can be viewed as acceptable clinical evidence by HTA agencies when these study designs are justifiable and when treatment effect can be convincingly demonstrated, but their use is currently limited.

Keywords: clinical effectiveness, decision making, evidence-based medicine, health technology assessment.

Introduction

Health technology assessment (HTA) can be broadly defined as the evaluation of health care interventions in the context of their implications to the wider health system. HTA aims to systematically assess the clinical value (i.e., comparative health benefits) and the economic value (i.e., value for money) of interventions to inform decisions regarding their reimbursement and uptake [1].

To assess the clinical value of a new intervention, it is necessary to compare it against currently available interventions in terms of patient-relevant outcomes such as efficacy, safety, and health-related quality of life. To ensure accurate and objective comparisons between interventions, the best available clinical evidence must be used. Randomized controlled trials (RCTs) have historically been considered the gold standard in the hierarchy of clinical evidence, surpassed only by meta-analyses of RCTs, whereas nonrandomized studies and uncontrolled studies are considered weaker evidence (Fig. 1) [2]. Compared with non-RCTs, RCTs minimize the likelihood of confounding factors influencing the results and therefore produce a more robust and less biased estimation of treatment effect. For this reason, HTA agencies usually express a preference for RCT evidence to assess comparative effectiveness [3–5].

Nevertheless, there are situations in which it is not ethical, feasible, or practical to conduct an RCT [6,7]. For example, it may be unethical to offer a placebo or an intervention that is hypothesized to be less effective than the intervention under evaluation (e.g., in immediately life-threatening disorders). Alternatively, the disease may be so rare that it would be difficult to recruit enough participants to detect statistically significant differences between treatment arms (e.g., rare genetic disorders), or there simply may be no established treatment options to compare against (e.g., some advanced cancers). In such cases, noncomparative studies may provide the "best available"

* *Address correspondence to:* Elizabeth A. Griffiths, PAREXEL Access Consulting, PAREXEL International, 160 Euston Road, London NW12DX, UK.

E-mail: egriffiths27@gmail.com.

1098-3015$36.00 – see front matter

http://dx.doi.org/10.1016/j.jval.2017.06.015

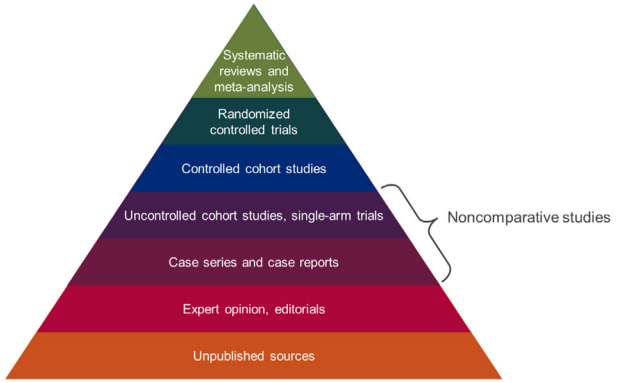

Fig. 1 – Evidence hierarchy of clinical studies. Adapted from Akobeng [2].

evidence. In the clinical trial setting, noncomparative studies encompass a range of designs including dose-ranging studies, single-arm trials, case series, and case reports [8–11]. In the "real-world" setting (i.e., outside of the typical clinical trial setting), this may include registry studies, claims data, and some observational designs [12]. Although considered less robust than RCTs, noncomparative studies can—and do—inform health care decision making.

In the era of biomarker-based, "personalized" medicine and conditional regulatory approvals based on immature clinical data [13], there may be cases in which reimbursement decisions will need to be made on the basis of noncomparative studies such as phase 1 or phase 2 trials (e.g., dose-ranging or single-arm trials), with this trend anticipated to continue [14,15]. Both the US Food and Drug Administration (FDA) and the European Medicines Agency (EMA) have developed mechanisms to facilitate earlier patient access to promising medicines, such as breakthrough status (FDA), the accelerated approval pathway (FDA), and the adaptive pathways pilot (EMA). For example, the FDA approved ceritinib for non–small cell lung cancer and pembrolizumab for melanoma in 2014 under both the FDA breakthrough status and accelerated approval processes, supported by only phase 1 data [16,17]. Although these expedited regulatory pathways can ensure earlier market authorization, to achieve patient access, public reimbursement must also be achieved, which typically requires previous recommendation by an HTA body.

It will therefore be important to understand how HTA bodies react to noncomparative evidence, and whether a positive outcome is achievable with such data. This study aimed to address the following research questions to characterize the role of noncomparative evidence in HTA decision making, which may help inform future HTA submissions:

1. What are HTA agencies' perceptions of noncomparative evidence, on the basis of recommendation and nonrecommendation rates, and what are the key differences in these rates between agencies?
2. Do recommendation rates for submissions presenting noncomparative evidence differ from rates for submissions presenting RCT evidence?
3. Are there any differences in recommendation rates/perceptions by parameters such as disease area and size of the patient population?

For this research, three HTA bodies were selected: the National Institute for Health and Care Excellence (NICE) in the United Kingdom, the Canadian Agency for Drugs and Technologies in Health (CADTH) in Canada, and the Institute for Quality and Efficiency in Health Care (*Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen* [IQWiG]) in Germany. These agencies were chosen because they represent key jurisdictions that use varying criteria to inform decision making in different settings, and release publicly available and transparent appraisal documents for all interventions that are evaluated.

Methods

Data Sources

The Web sites of the three jurisdictions—NICE (https://www.nice.org.uk/), CADTH (https://www.cadth.ca/), and IQWiG (https://www.iqwig.de/)—were systematically searched for publicly available HTA appraisals published between January 2010 and December 2015. This date range was chosen because it allowed for a wide range of appraisals across the three agencies to be analyzed and it reflects recent decision-making trends; the evolution of HTA processes is such that decisions published before 2010 may be less relevant to today's reimbursement landscape. The IQWiG and the CADTH pan-Canadian Oncology Drug Review (pCODR) did not publish their first decisions until 2011 and 2012, respectively; in addition, all three agencies revisit their methods every few years.

Appraisal Selection

The inclusion criteria encompassed all single technology appraisals for pharmaceutical interventions, irrespective of indication or outcome. Multiple technology appraisals, appraisals for vaccines and devices, requests for advice, and health economic dossiers were excluded.

Data Extraction

For each appraisal meeting, the inclusion criteria, indication, date issued, outcome, and clinical evidence presented were extracted. Appraisals were classified as recommended, restricted, or not recommended. Detailed definitions of these outcomes are described in Appendix Table 1 in Supplemental Materials found at http://dx.doi.org/10.1016/j.jval.2017.06.015.

The clinical evidence was defined as the data presented to support the clinical case, for efficacy, safety, or quality-of-life outcomes, as categorized by the HTA agency. Data used to inform economic modeling were not included. The clinical evidence was first categorized into two categories, comparative study or noncomparative study, using the following definitions:

1. *Comparative study*: Any study against an active or placebo comparator:
   - Active-controlled RCT: randomized controlled study against a relevant active comparator (note that best supportive care was considered to fall under this category);
   - Placebo-controlled RCT: randomized controlled study against a placebo comparator (including vehicle-controlled studies, sham injections, and studies in which the active intervention + X was compared against placebo + X);
   - Other comparative study: study designs encompassing nonrandomized, controlled trials.
2. *Noncomparative study*: Any study that did not compare against an active comparator or placebo. This included, but was not limited to, single-arm trials, dose-ranging studies, registry studies, compassionate use programs, and uncontrolled extension studies.

Noncomparative studies were further classified into single-arm trials, single-arm extensions of comparative trials, dose-ranging studies, and other (see Appendix Table 2 in Supplemental