The Need to Consider the Impact of Drivers of Relative Treatment Effects Beyond Treatment Choice in Cost-Effectiveness Analyses

Stevens W[1], Krackow L[1], Neslusan C[2]

[1]Medicus Economics, LLC, Milton, MA, USA; [2]Johnson & Johnson, Titusville, NJ, USA

Poster Code: EE206

Background

- When assessing the value of new technologies, stakeholders often consider estimates of cost-effectiveness that use methods focused on the question of what value is added by an innovation in isolation. Factors such as where, how, and when the technology will be used are rarely fully incorporated into these calculations. This is despite multiple examples of efficacy data from clinical trials not being replicated in terms of measured effectiveness in retrospective analyses. Studies have documented cases in which real-world treatment effectiveness is greater than that suggested by the efficacy estimates obtained from trials [1], as well as cases in which the efficacy estimates obtained from trials are greater than real-world treatment effectiveness [2]
  - **Figure 1A** illustrates the traditional economic modeling assumption that efficacy is the best point estimate of effectiveness, which requires the acceptance of ceteris paribus (i.e., all else being equal) and is unlikely to hold in practice. **Figure 1B** illustrates the hypothetical impact of non-treatment drivers on absolute treatment effectiveness. The real-world effectiveness of Treatment (Tx) A is greater than the absolute efficacy of Tx A as a result of contextual factors x and y. For Tx B, these same factors result in real-world effectiveness of Tx B being less than the absolute efficacy of Tx B
- In other industries, technologies are routinely assessed "in place" to ascertain their value. For example, when we picture the potential impact of introducing artificial intelligence in healthcare, contextual factors such as where (e.g., primary care, hospitals, or specialist centers), how (e.g., decision automation or as a tool to inform physicians), and when (e.g., at diagnosis, prognosis, or treatment choice) this innovation will be used are key considerations [3]

Figure 1. The Role of Contextual Drivers in Determining Relative Effectiveness

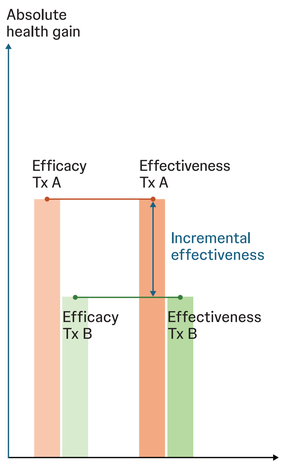


Figure 1A. Traditional CEA assumption of equality of efficacy and effectiveness

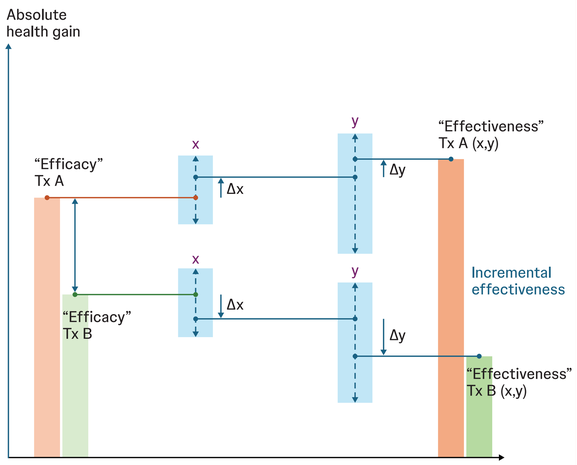


Figure 1B. A hypothetical example of how two different non-treatment drivers (x and y) can impact real-world effectiveness

CEA, cost-effectiveness analysis; Tx, treatment.

Research objective

- **To examine the extent to which contextual factors (i.e., non-treatment drivers of effectiveness) can contribute to variation in relative treatment effectiveness (i.e., real-world performance) as compared with the impact of relative efficacy**

Methods

- To illustrate the potential for non-treatment drivers to impact variation in real-world effectiveness, we selected major depressive disorder (MDD) for analysis as this therapeutic area is known for heterogeneity in patient presentation, treatment patterns, and outcomes
  - We reviewed real-world studies from 2005 to 2025 that examined non-treatment drivers of real-world effectiveness
  - Studies were selected for consideration if the effect size of the non-treatment driver showed statistical significance at the 5% level (P <0.05)
  - Variation in the estimates of relative treatment effectiveness resulting from the impact of non-treatment drivers was sourced directly from the identified studies (**Table 1**)
    - The odds ratio of relative effectiveness associated with the existence or non-existence of non-treatment drivers was calculated
  - Estimates were benchmarked to a relative efficacy estimate obtained from a highly cited systematic review and network meta-analysis that evaluated 21 different pharmaceuticals for the treatment of MDD in adults [4]. This study included 522 randomized controlled trials (RCTs); only RCTs in which >20% of the population had a co-existing psychiatric comorbidity were excluded
    - The odds ratio of relative efficacy between the most efficacious treatment (amitriptyline) and the least efficacious treatment (reboxetine) was used to reflect the widest degree of variation in outcomes between treatments

Results

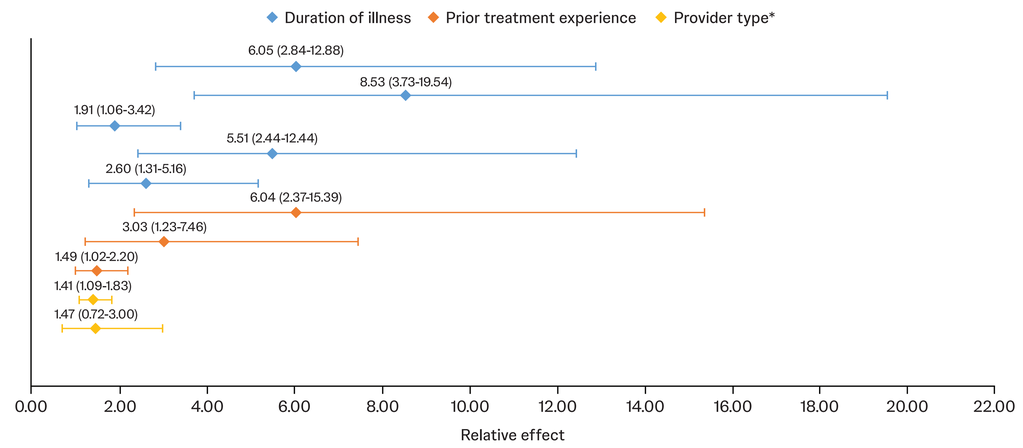
- The review of MDD-specific real-world studies yielded empirical evidence on three non-treatment drivers of effectiveness:
  1. Duration of untreated illness: four studies estimated the effect size and variation in relative effectiveness of treatment associated with the duration of illness prior to treatment initiation
  2. Prior treatment experience: three studies estimated the effect size and variation in relative effectiveness of treatment associated with prior experience of drug therapy for MDD before treatment initiation
  3. Type of provider: one study estimated the effect size and variation in relative effectiveness of treatment associated with whether the prescriber was a primary care physician or psychiatrist. Two estimates with associated variations were reported, one for response to treatment and the second for remission of MDD
- **Table 1** presents the studies included in the review, outcomes evaluated, directional impact on effectiveness, and the calculated odds ratios and associated confidence intervals (CIs)
  - Calculated odds ratios ranged from 1.91 to 8.53 for the studies that measured the impact of duration of untreated illness, from 1.49 to 6.04 for the studies that measured non-presence versus presence of prior drug therapy, and from 1.41 to 1.47 for the study that evaluated the impact of prescription by psychiatrist versus a primary care prescriber (**Table 1** and **Figure 2**)
  - In comparison, the odds ratio of efficacy between the most efficacious treatment (amitriptyline) and the least efficacious treatment (reboxetine) was 1.50 (95% CI, 1.07-2.07)

Table 1. Impact of Non-Treatment Drivers on Treatment Effectiveness

Study	Driver	Impact on effectiveness	Calculated ORs (95% CI)
Okuda et al. [5]	Duration of illness	Remission rate at Week 8 was lower for the group with longer duration of illness than the group with shorter duration of illness	6.05 (2.84-12.88)
Okuda et al. [5]	Duration of illness	Response rate at Week 8 was lower for the group with longer duration of illness than the group with shorter duration of illness	8.53 (3.73-19.54)
Bukh et al. [6]	Duration of illness	Remission rate at Week 12 was lower for the group with longer duration of illness than the group with shorter duration of illness	1.91 (1.06-3.42)
Ghio et al. [7]	Duration of illness	Remission rate at Week 24 was lower for the group with longer duration of illness than the group with shorter duration of illness	5.51 (2.44-12.44)
de Diego-Adelino et al. [8]	Duration of illness	Mean time to sustained response was greater for the group with longer duration of illness than the group with shorter duration of illness	2.60 (1.31-5.16)
Blumberger et al. [9]	Prior treatment experience	Remission rate was lower for treatment-experienced patients than those who were treatment naïve	6.04 (2.37-15.39)
Tew et al. [10]	Prior treatment experience	Remission rate was lower for treatment-experienced patients than those who were treatment naïve	3.03 (1.23-7.46)
Hsu et al. [11]	Prior treatment experience	Remission rate was lower for treatment-experienced patients than those who were treatment naïve	1.49 (1.02-2.20)
Cuijpers et al. [12]	Provider type*	Response rate was higher for those whose treatment was prescribed by a psychiatrist than those whose treatment was prescribed by a primary care physician	1.41 (1.09-1.83)
Cuijpers et al. [12]	Provider type*	Remission rate was higher for those whose treatment was prescribed by a psychiatrist than those whose treatment was prescribed by a primary care physician	1.47 (0.72-3.00)

CI, confidence interval; OR, odds ratio.
*For this driver, relative risk ratios were calculated, not ORs, specifically due to data limitations.

Figure 2. ORs Measuring Relative Effectiveness as a Function of Non-Treatment Drivers



OR, odds ratio.
*For this driver, relative risk ratios were calculated, not ORs, specifically due to data limitations.

Discussion

- The review of the literature in MDD identified empirical evidence for three determinants of variation in treatment effectiveness beyond relative treatment efficacy: duration of untreated illness, prior treatment experience, and provider type
- Notably, in some cases, the estimate of the impact of a non-treatment driver on the variation in real-world effectiveness was substantially greater than a benchmark estimate of marginal effectiveness based on the choice of treatment alone
- Relying solely on RCT estimates of efficacy for economic analyses can be problematic in multiple ways [13]. For instance, the relative effectiveness of any new technology is a time-variant variable; it may change as practice and systems evolve. In addition, RCT populations and treatment pathways may not be fully reflective of real-world patient populations or treatment pathways [14,15]
- Factors beyond the choice of treatment can have opposing effects on the effectiveness of technologies. For example, in MDD there is empirical evidence that older patients, even independent of physical condition and comorbidity, may respond more slowly to antidepressants [16], and RCTs often specify an upper age-limit exclusion criterion. As such, treatment effectiveness in this population would be expected to be less than efficacy estimates would suggest [17]. Conversely, MDD RCTs tend to include more treatment-experienced patients than seen in real-world settings [18], and these patients have been shown to have lower response rates compared with those without prior exposure [19]. Ignoring this non-treatment driver would result in an underestimation of effectiveness
- Although the analysis included studies where effectiveness estimates were not fully adjusted for all possible factors, those identified are commonly thought to be important considerations for outcomes in MDD
- In addition, these findings are a function of what has been published on MDD and the studies in this review were not designed to comprehensively investigate the impact of contextual factors on patient outcomes
- Further empirical work is warranted to not only refine the scale of the effects reported, but also explore other possible contextual factors (e.g., variation in health system performance) and their interactions

Conclusions

- **Contextual factors can be critical drivers of the variation in real-world treatment effectiveness**
- **If resource-allocation decisions are informed by analyses that assume real-world outcomes are not affected by the heterogeneity of either patient experience or where, how, and when treatments are delivered, resources can be misallocated, which can result in non-optimal patient outcomes**
- **In translating efficacy into effectiveness (a critical step in assessing the value of technologies via economic modeling), researchers should evaluate the extent to which factors beyond the choice of treatment can impact real-world performance**

Acknowledgment

Editorial support was provided by Kim Caldwell, PhD, CMPP™, of Lumanity Communications Inc., and funded by Johnson & Johnson.

Disclosures

This study was sponsored by Johnson & Johnson. WS and LK are employees of Medicus Economics, LLC. CN is an employee of and holds stock in Johnson & Johnson.

The Professional Society for Health Economics and Outcomes Research (ISPOR); May 17-20, 2026; Philadelphia, PA, USA
Poster Session 2
Monday, May 18, 4:00 PM – 7:00 PM
Value in Health, Volume 29, Issue S6

REFERENCES

1. Stevens, W., et al., Health Econ, 2020. 29(5): p. 580-590. 2. Hunsley, J. and C.M. Lee, Prof Psychol Res Pr, 2007. 38(1): p. 21. 3. Alowais, S.A., et al., BMC Med Edu, 2023. 23(1): p. 689. 4. Cipriani, A., et al., Lancet, 2018. 391(10128): p. 1357-1366. 5. Okuda, A., et al., Psychiatry Clin Neurosci, 2010. 64(3): p. 268-273. 6. Bukh, J.D., et al., J Affect Disord, 2013. 145(1): p. 42-48. 7. Ghio, L., et al., J Affect Disord, 2015. 175: p. 224-228. 8. de Diego-Adelino, J., et al., J Affect Disord, 2010. 120(1-3): p. 221-225. 9. Blumberger, D.M., et al., J Psychiatr Psych Res, 2011. 45(7): p. 896-901. 10. Tew, J.D. Jr., et al., Am J Geriatric Psychiatry. 2006. 14(11): p. 957-965. 11. Hsu, J.H., et al., Am J Geriatric Psychiatry, 2016. 24(10): p. 918-922. 12. Cuijpers, P., et al., J Affect Dis, 2023. 339: p. 660-675. 13. Deaton, A. and N. Cartwright, Soc Sci Med, 2018. 210: p. 2-21. 14. Birch, S. and A. Gafni, J Health Econ, 1993. 12(4): p. 469-476. 15. Tan, Y.Y., et al., Lancet Healthy Longev, 2022. 3(10): p. e674-e689. 16. Mandelli, L., et al., Psychiatry Res, 2007. 152(1): p. 37-44. 17. Giron, M.S., J. Fastbom, and B. Winblad, Int J Geriatr Psychiatry, 2005. 20(3): p. 201-217. 18. Zetin, M., C.T. Hoepner, and L. Bjornson, Psychopharmacol Bull, 2006. 39(1): p. 38. 19. Pigott, H.E., et al., Psychother Psychosom, 2010. 79(5): p. 267-279.