

# Let's all get on the same page: A commentary on “Defining response and nonresponse to PTSD treatments: A systematic review”

Sadie E. Larsen<sup>1,2</sup>  | Lauren M. Sippel<sup>3,4</sup> | Paula P. Schnurr<sup>3,4</sup>

<sup>1</sup>Milwaukee VA Medical Center, Milwaukee, WI, USA

<sup>2</sup>Medical College of Wisconsin, Wauwatosa, WI, USA

<sup>3</sup>National Center for Posttraumatic Stress Disorder, White River Junction, VT, USA

<sup>4</sup>Geisel School of Medicine at Dartmouth, Hanover, NH, USA

**Correspondence:** Sadie E. Larsen, Milwaukee VA Medical Center, 5000 W. National Ave, Milwaukee, WI 53296, USA.

Email: Sadie.Larsen@va.gov

## 1 | INTRODUCTION

Varker et al. (2020) present a good synthesis of a literature that defies easy summarization. The field of mental health is centrally focused on developing treatments that improve patients' lives by reducing symptoms, improving functioning, and enhancing quality of life. Given these goals, the findings of this systematic review—namely, that many posttraumatic stress (PTSD) treatment studies (40% post-2010) do not operationalize treatment response at all and those that do have vastly different definitions—are concerning. Below we comment on Varker et al.'s (2020) stated concerns with the literature and their recommendations for validating categorical definitions of treatment outcomes. We then propose research strategies that would advance knowledge about optimal definitions of treatment response, nonresponse, and resistance.

To set the stage, we assert that these definitions and distinctions are important because providers want to be able to tell patients what to expect from treatment, including the likelihood that they will improve, and by how much. With the current state of the literature, in which definitions of response and nonresponse are so heterogeneously defined, it is difficult to meaningfully summarize and compare outcomes. For instance, consider three studies that assess PTSD with the Clinician-Administered PTSD Scale (CAPS-IV) for the Diagnostic and Statistical Manual of Mental Disorders (4th edition; DSM). One defines response as a 10-point reduction

on the CAPS-IV, another defines response as a 15-point reduction, and a third defines only remission as a CAPS-IV score of <20. How can the studies be compared? Another study uses the CAPS for 5th edition of the DSM (CAPS-5) and reports that, on average, there was a statistically significant effect size of 1.0 for the treatment. In addition to using a different scale, this study's effect size offers no information about the effects on individual patients. Did some show more improvement and others not at all? Or did everyone equally improve?

As these examples highlight, categorical definitions of treatment outcome can be distinguished by whether they focus on *change*, such as response, or on *endpoints*, such as remission. Definitions of change include absolute and relative measures. Absolute measures of change include a fixed magnitude of change on a scale, such as 10 points on the CAPS-IV. Relative change can be defined with respect to an individual patient (e.g., 50% decrease from baseline) or a scale (e.g., 20% of the total possible range of the scale; discussed below). Defining response in absolute terms is simple and easy to understand, but is specific to the measure used and is invariant across baseline severity, whereas defining response in relative terms is helpful for making comparisons across different measures. Definitions of endpoints are typically measure-specific, such as remission defined as a score of <20 on the CAPS-IV. Change and endpoints are sometimes combined, such as loss of diagnosis defined as a combination

of response (CAPS-IV decrease of  $\geq 10$  points), no longer meeting symptom diagnostic criteria, and a post-treatment CAPS-IV score  $< 45$  (Schnurr & Lunney, 2016).

Randomized clinical trials most often report results in terms of whether the treatment led to a statistically significantly greater effect size change in symptoms than a comparison condition. Effect size is an important metric of *change* that can be readily compared across treatments and across different measures. However, an effect size does not convey fundamental information that patients will want: the likelihood of improvement or recovery. What are the chances that this treatment, on average, helps people to make a change that is noticeable in terms of symptoms and improvements in quality of life and functioning? An effect size is not helpful for conveying this information to an individual patient; rather, we need the type of categorical response categories highlighted in this systematic review.

Thus, we firmly reinforce Varder et al.'s call for standardizing definitions of treatment response, remission, and nonresponse, which would have at least two clear consequences. First, it would be more feasible to directly compare outcomes across studies. Second, it would improve shared decision-making and communication with patients. For example, telling a patient that they have a 50% likelihood of no longer having a PTSD diagnosis after a given treatment is more interpretable and meaningful than telling them that the effect size of the treatment is 1.0. Having standardized definitions can also guide clinicians in deciding how long treatment should continue and/or whether a new approach should be considered—that is, if the patient has improved but still meets diagnostic criteria for PTSD or has significant symptoms, should some sessions be added or a new treatment approach considered?

## 2 | AUTHOR CONCERNS

Varker et al. (2020) raised several concerns about the state of the evidence they reviewed. First, they questioned the empirical basis for most categorical definitions of treatment outcome. We completely agree that empirical validation of the definitions of response for various measures is needed. For instance, within the depression literature, “response” is differentiated from “remission,” partially because remitters have much better outcomes in terms of functioning, prognosis, and stability (Rush et al., 2006). One change metric that has been validated for PTSD is a 10-point reduction on the CAPS-IV as a measure of response (Schnurr & Lunney, 2016). This was done by benchmarking definitions of response, loss of diagnosis, and remission with improvement and good endpoints on various measures of functioning and quality of life (Schnurr & Lunney, 2016). Other measures, including the CAPS-5, could benefit from such an approach.

Varker et al.'s (2020) second concern relates to absolute reductions in symptom severity, not taking into account baseline severity, so that a person who still had significant symptoms could be classified as having responded to treatment. We agree that anyone with significant residual symptoms should not be considered a treatment success, but disagree that severity should be incorporated into the definition of response. It is important to capture improvement, even if a person is still symptomatic. Kazdin (2001) gives the example of a patient with agoraphobia who improves just enough to accompany a spouse to social events, a change that may be quite meaningful within the relationship. We encourage more widespread reporting of both change (response) and endpoint (remission) outcomes, as they convey fundamentally different information and help to paint a more complete picture of patient recovery. One other option, which is acknowledged by the authors, is to utilize a measure that incorporates both change and endpoint, such as the loss of the diagnosis operationalization mentioned above (Schnurr & Lunney, 2016).

Varker et al.'s (2020) third concern relates to the absence of information about other meaningful outcomes (e.g., functioning) in definitions of treatment response. We do not share this concern. Indeed, it is common in the mental health treatment outcome literature to base definitions of response and remission on the primary symptoms being treated, partially because functioning can be affected by comorbid conditions and is thus a nonspecific measure of change (Rush et al., 2006).

The authors' final concern relates to a failure of investigators to specify how long symptoms have to be absent to meet definitions of remission, noting that this could confuse definitions such as remission and loss of diagnosis. We agree that it is important to clarify the timeframe for various measures.

## 3 | VARDER ET AL.'S RECOMMENDATIONS

Varker et al. (2020) recommend defining treatment response as a reduction of baseline symptomatology of 30%–50%, a clinically meaningful response on the CAPS-IV as 15 points, and including other symptom or functioning measures in more rigorous definitions of treatment response. Given the frequently used definition of response in the depression literature as 50% improvement (Rush et al., 2006), defining response as a 50% change might be useful to consider. However, there is an inherent problem with using percent change from baseline to define response: The amount of change needed to meet a threshold is higher for more severe individuals. To illustrate, Varder et al. proposed 15 points on the CAPS-IV would be only 21% improvement in a patient with a baseline score of 70, but 30% improvement in a patient with a baseline of 50.

We recognize that an individual who improved from 70 to 55 on the CAPS-IV still has significant symptoms, but believe that response is response, and should not be conditional on the absolute values of symptoms before treatment. Instead, additional endpoints, such as remission, or a composite loss of diagnosis measure (Schnurr & Lunney, 2016) is a better way to capture change in post-treatment severity. An alternative to percent change from baseline is reporting change as a percent of a scale's range. For example, 15 points on the CAPS-IV is 11% of the scale's range (0–136). Translated to the CAPS-5 (range 0–80), 11% would be 9 points.

Regarding the authors' suggestion to use measures of functioning in rigorous definitions of response, we suggest an alternative strategy—to benchmark symptom-based definitions using these kinds of measures (Schnurr & Lunney, 2016). Additionally, it could be informative for PTSD treatment trials to report categorical response measures for comorbid symptoms and functioning, as well as for PTSD. These measures convey meaningfully different information and would inform models of psychopathology, including their shared versus unique qualities and mechanisms.

Vardar et al. next recommend defining *remission* as 20 points or less on the CAPS-IV (the previously used definition for remission on the CAPS-IV was a score <20, not  $\leq 20$ ). Given the change in CAPS-5 scaling to 0–80, we instead propose that the definition be translated to a score of CAPS-5 <12 (based on the percentage of range). However, it is important to consider definitions of remission for other measures of PTSD as outlined in the DSM-5 and to align definitions across measures. It may be useful to operationalize response and remission based on frequently used scales such as the CAPS-5, but empirically validated crosswalks are needed to permit optimal comparisons across measures. The goal in any such definition is to capture a state in which symptoms are absent or at a mild, subclinical level.

After proposing this definition of remission ( $\leq 20$  on CAPS-IV), the authors then recommend defining *recovery* as maintaining remission for at least 6 months after treatment completion. We propose that (a) remission at the end of a treatment episode is an important outcome (and the most commonly available timepoint) and should be preserved, and (b) it would be useful to report on remission that is *sustained over time* (note that sustained remission is labeled “recovery” in the depression literature; Rush et al., 2006). We suggest that, for PTSD treatment trials, sustained remission be distinguished from *recovery* as defined more holistically by organizations such as the Substance Abuse and Mental Health Services Administration, which typically includes broader constructs such as living a self-directed life, having a sense of purpose, and participating in one's community (Sheedy & Whitter, 2009).

Finally, Varker et al. (2020) recommend distinguishing treatment nonresponse definitions (i.e., nonresponse vs. worsening vs. treatment resistance). We agree that these distinctions are important. In particular, treatment resistance is a different—although related—construct. Whereas response/nonresponse is a treatment outcome and is needed to define resistance, treatment resistance also needs to be defined based on the type of treatment received as well as history of response/nonresponse to multiple treatments. Therefore, treatment resistance could not be defined as an outcome in a study that did not administer a sequence of treatments or carefully assess treatment history to determine type and adequacy of prior treatment.

## 4 | A PROPOSED RESEARCH AGENDA

Several recommendations for future research flow from our own concerns and recommendations for defining response and nonresponse. First, we recommend validating definitions of treatment response and remission against measures that are meaningful to patients. This could involve benchmarking PTSD symptom measures against measures of functioning or quality of life (Schnurr & Lunney, 2016). Currently, as Vardar et al. describe, the field is hampered by varied definitions—but also by very little empirical basis to argue for one standard definition or another. This type of research would represent one step toward ensuring that definitions of symptom change reflect meaningful outcomes.

Second, defining meaningful change in treatment would be enriched by attention to patient perspectives. Even inclusion of measures of functioning is not a cure for the difficulty in defining meaningful response. As others have argued, patients' own definitions of what entails meaningful change and endpoints will likely only somewhat overlap with existing symptoms and functioning measures (Kazdin, 2001). In weighing treatment options and determining the end of a treatment episode, patients' experiences and preferences must be considered. In doing so, we may be surprised. For example, Zimmerman (2012) described studies of patients treated for depression, finding that some patients who remained symptomatic per a standardized depression measure considered themselves to be in remission, and others who were asymptomatic per the same measure considered themselves in need of further treatment for depression.

Finally, it would be useful to validate a crosswalk of various commonly used symptom measures to ensure that researchers can meaningfully compare results from studies that used different scales. For example, if response were defined as a given percentage of a scale's range, the definition could be applied across scales and then validated

to determine if the definition is comparable across scales (i.e., that 20% on Scale A is comparable to 20% on Scale B). This would be preferable to using a relative change measure such as percent change from baseline, for the reasons noted above.

## 5 | CONCLUSION

Almost 20 years ago, in a commentary on a review of literature on clinically significant change in treatments for conduct disorder, Kazdin (2001) discussed the challenges in operationalizing clinically significant change. Unfortunately, the concerns he raised still apply broadly to the treatment outcome literature in PTSD and other mental health disorders, although we believe progress is being made. The review by Varker et al. (2020) adds to that progress. The authors are to be commended for synthesizing a difficult literature and for proposing ways to enhance the measurement of treatment outcomes. Although we agree with their call for standardization of treatment response definitions, we argue for a different approach—one that maximizes the types of information conveyed through the use of multiple measures and several ways of categorizing changes on those measures. Further, we recommend that researchers examine outcome data in ways that will allow these data to inform the choices the field makes about how to define different categorical outcomes. Underlying our approach is our belief in the importance of ensuring that these definitions translate into meaningful changes in patients' lives or as Kazdin (2001) said, "when a difference makes a real difference" (p. 461).

## ACKNOWLEDGMENTS

This manuscript is partially the result of work supported with resources and the use of facilities at the Milwaukee VA. The contents do not represent the views of the U.S. Department of Veterans Affairs or the United States Government.

## ORCID

Sadie E. Larsen  <https://orcid.org/0000-0001-7365-6199>

## REFERENCES

- Kazdin, A. E. (2001). Almost clinically significant ( $p < .10$ ): Current measures may only approach clinical significance. *Clinical Psychology: Science and Practice*, 8(4), 455–462. <https://doi.org/10.1093/clipsy.8.4.455>
- Rush, A. J., Kraemer, H. C., Sackeim, H. A., Fava, M., Trivedi, M. H., Frank, E., ... ACNP Task Force. (2006). Report by the ACNP Task Force on response and remission in major depressive disorder. *Neuropsychopharmacology*, 31(9), 1841–1853. <https://doi.org/10.1038/sj.npp.1301131>
- Schnurr, P. P., & Lunney, C. A. (2016). Symptom benchmarks of improved quality of life in PTSD. *Depression and Anxiety*, 33(3), 247–255. <https://doi.org/10.1002/da.22477>
- Sheedy, C. K., & Whitter, M. (2009). *Guiding principles and elements of recovery-oriented systems of care: What do we know from the research?* Center for Substance Abuse Treatment, Substance Abuse and Mental Health Services Administration. Retrieved from [https://www.samhsa.gov/sites/default/files/partnersforrecovery/docs/Guiding\\_Principles\\_Whitepaper.pdf](https://www.samhsa.gov/sites/default/files/partnersforrecovery/docs/Guiding_Principles_Whitepaper.pdf)
- Varker, T., Kartal, D., Watson, L., Freijah, I., O'Donnell, M., Forbes, D., ... Hinton, M. (2020). Defining response and non-response to PTSD treatments: A systematic review. *Clinical Psychology: Science and Practice*, 27(4), e12355. <https://doi.org/10.1111/cpsp.12355>
- Zimmerman, M. (2012). Discordance between researchers and patients in defining remission from depression. *Journal of Clinical Psychiatry*, 73, 1262–1263. <https://doi.org/10.4088/JCP.12ac08081>