Received: 12 October 2022 | Revised: 30 December 2022 | Accepted: 8 January 2023

DOI: 10.1002/lio2.1016

ORIGINAL RESEARCH

Laryngoscope Investigative Otolaryngology

Patient reported outcome measures: The impact of environment on VHI-10 responses

Eliezer C. Kinberg MD[1] | Amir Taree MD[2] | Mingyang Gray MD, MPH[1] | Peter Filip MD[1] | Mark Courey MD[1]

[1]Department of Otolaryngology-Head and Neck Surgery, Icahn School of Medicine at Mount Sinai, New York, New York, USA

[2]Icahn School of Medicine at Mount Sinai, New York, New York, USA

Correspondence
Eliezer C. Kinberg, Icahn School of Medicine at Mount Sinai Department of Otolaryngology-Head and Neck Surgery, 1 Gustave L Levy Place, New York, NY 10029, USA.
Email: eli.kinberg@gmail.com

Abstract

Objective: A key outcome measure in the clinical evaluation of dysphonia is the Voice Handicap Index (VHI-10). The clinical validity of the VHI-10 was established from surveys administered in the physician's office. We aim to understand whether VHI-10 responses remain reliable when the questionnaire is completed in settings other than the physician's office.

Methods: This is a prospective observational study conducted over a 3-month period in the outpatient laryngology setting. Thirty-five adult patients presenting with a complaint of dysphonia, which was symptomatically stable for the preceding 3 months, were identified. Each patient completed a VHI-10 survey during the initial office visit, followed by three weekly out-of-office (termed "ambulatory") VHI-10 surveys, over the course of 12 weeks. The specific setting in which the patient completed the survey was recorded (social, home, or work). The Minimal Clinically Important Difference (MCID) is defined as 6 points based on existing literature. T-tests and a test of one proportion were used for analysis.

Results: A total of 553 responses were collected. Of these, 347 ambulatory scores (63%) differed from the Office score by at least the MCID. Specifically, 94 (27%) were higher than the in-office score by 6 or more points while 253 (73%) were lower.

Conclusion: The setting in which the VHI-10 is completed affects how the patient answers the questions. The score is dynamic, reflecting effects of the patients' environment during completion. Utilization of VHI-10 scores to measure clinical treatment response is only valid if each response is obtained in the same setting.

Level of Evidence: 4

KEYWORDS
outcomes/cost effectiveness, voice disorders, voice/dysphonia

This manuscript was presented as a poster presentation at the 2020 Annual Meeting of the American Academy of Otolaryngology.

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.
© 2023 The Authors. *Laryngoscope Investigative Otolaryngology* published by Wiley Periodicals LLC on behalf of The Triological Society.

1 | INTRODUCTION

Patient reported outcome measures (PROMs) are an important component of health care delivery and outcomes assessments.[1–3] In certain instances, the Center for Medicaid and Medicare Services even relies upon PROMs in determining hospital quality and reimbursement allocations.[4] In fields such as oncology, in which symptoms are relied upon to guide treatment decisions, the literature demonstrates that ambulatory PROMs provide valuable data which effectively contributes to clinical decision making.[3,5] Nevertheless, there remains a gap in the literature evaluating whether the setting in which the PROM is completed impacts its results.

Within otolaryngology, the Voice Handicap Index 10 (VHI-10) is one example of a commonly used patient reported outcomes measure. The VHI-10 is a scale which was designed to evaluate the impact of voice disorders on patients' quality of life by providing a subjective self-assessment of voice.[6] The scale quantifies a patient's own perception of handicap, disability and distress attributed by his/her voice.[7] The validation studies of the VHI-10, including the initial proof of concept and subsequent assessments of its Minimal Clinical Important Difference (MCID), were conducted based on surveys collected in the otolaryngologists' office.[8,9]

In considering the clinical applicability of PROMs, it is important to ensure that the PROM is both validated and correctly measures its intended outcome. For instance, if a PROM designed to measure voice usage was administered to a post-operative cohort during a period of prescribed voice rest, one can envision that the results might be statistically significant yet not translatable to a healthy patient cohort. This element of PROM validation is difficult to quantify and remains a challenge to their usage.[10]

To date, the literature has not described the impact of environment on PROMs. Specifically, some PROMs may establish their validity from responses collected in the waiting room of a physician's office while others are computer-based, and therefore, validated from responses collected in patients' home or workplace. Subsequently, the PROM is considered statistically validated and it is assumed that the PROM provides an accurate measure whether it is then collected in a physician's office or patient's home. The question of whether a PROM validated in one setting can then be relied upon in another setting has not been addressed in the literature.

In the COVID-19 era, which has accelerated the integration of telemedicine into the daily clinic routine, it is important to understand if PROMs collected in non-office settings provide equivalent information to PROMs collected in the physicians' office. Specific to voice, it stands to reason that patients' responses to a VHI-10 survey in the office might be different than those provided in other settings. Hypothetically, while in the otolaryngology office, a patient may be focused on their voice complaint and indicate a worse dysphonia than they might perceive in another setting such as their home or workplace. In an alternative situation, a patient may be sitting quietly filling out forms in the waiting room and not using their voice. They may, therefore, be less bothered by their dysphonia in that moment and provide a lower VHI-10 than they might in a heavy voice-use situation such as their work. As a validated measure, these factors are presumed to average out when a VHI-10 score is collected in the in-office setting. As such, in-office VHI-10 scores are considered representative of a patients' generalized perception of their dysphonia.

It is unknown whether VHI-10 scores collected in the out-of-office, termed "ambulatory," setting are similarly broadly representative of a patients' overall dysphonia burden. To that end, we hypothesize that patients' perception of their voice will vary depending upon the location and situation in which they complete the VHI-10 form. The purpose of this study was, therefore, to fill this gap in the literature by assessing differences in patient responses to the VHI-10 survey in the home, work and social setting as compared to the physicians' office. Understanding anticipated differences in VHI-10 scores collected in in-office versus ambulatory settings will help clinicians correctly interpret the data and guide patient care appropriately. If a particular environment carries undue impact on the PROM result, clinicians will be able to understand the PROM in a more nuanced light and interpret clinical outcomes more accurately.

2 | MATERIALS AND METHODS

This is a prospective observational study conducted at a tertiary medical center. IRB approval was obtained from Mount Sinai Hospital. Following IRB approval, adult patients presenting to the laryngology office with a self-reported 3-month history of stable dysphonia were identified. Patients were excluded if they had a history of laryngeal carcinoma or treatment with head and neck radiation. While other disease states were not intentionally excluded, the patients in our sample carried diagnoses of muscle tension dysphonia, vocal fold atrophy, or vocal fold hypomobility. Patients were further excluded if they chose to undergo treatment for their dysphonia during the study period, either behavioral, medical, or surgical. Thirty-five patients who met the inclusion criteria consented to participate in the research study.

Each subject completed a VHI-10 survey during the initial office visit. Patients then completed three weekly ambulatory VHI-10 surveys, administered via e-mail utilizing the web-based REDcap survey tool, over the course of 12 weeks. These were e-mailed to patients via an automated system three times per week and could be completed on a mobile device or a computer. Typical VHI-10 instructions were provided. In addition to the VHI-10, patients were asked to identify if they were completing the survey while at home, at work, or in a social setting. The social setting was defined as any location other than home or work. We term these out-of-office scores "ambulatory." Patients were not prompted to respond to each survey; however, if a patient skipped 1 week of surveys, a reminder e-mail was sent out.

A paired *t*-test was utilized to analyze the office baseline score as compared to the aggregate ambulatory score. A test of one-proportion was utilized for analysis of sub-group results. Statistical significance was set at $p < .05$. The minimal clinically important difference (MCID), which is a measure that is utilized to provide clinical significance to outcomes measures that are otherwise based purely on

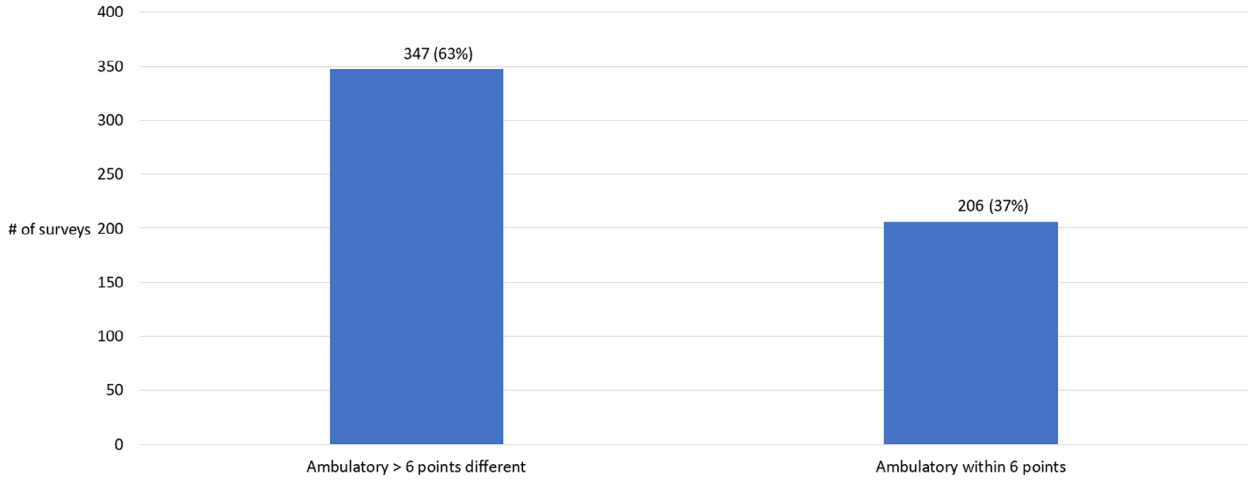

FIGURE 1 Total number of ambulatory versus in-office surveys meeting the MCID.

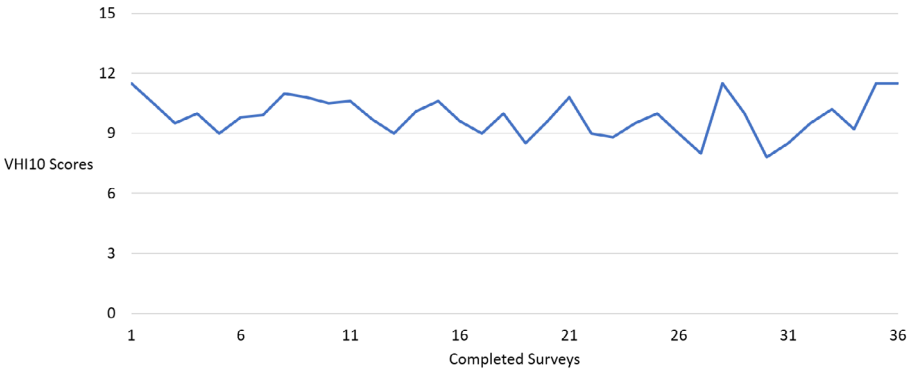

FIGURE 2 Stability of responses (*y* axis) across the 36 completed studies (*x* axis).

statistical significance, was defined as 6 points based on existing validation studies.[8]

When analyzing longitudinal data, one concern is the presence of "drift" in which changes to the data generation process impact the result being collected. Explained a different way, it is possible that subjects' responses to a PROM would change over the course of time simply because they have completed the form on multiple previous occasions. To account for the possibility of drift, we then analyzed the responses over time of the 10 patients within the population who completed all 36 surveys.

3 | RESULTS

A total of 553 survey responses were collected. There were 15 male and 20 female respondents with a mean age of 58 years. In the aggregate, the mean office VHI was 14 while the mean ambulatory VHI was 10.8. There were 347 ambulatory scores that differed from the in-office score by at least the minimal clinically important difference of 6 points, which is 63% of the total number of collected scores. Of these, 94 (27%) were higher than the in-office score by the MCID of 6 while 253 (73%) were lower ($p < .001$) (Figure 1).

Of the 10 patients who completed all 36 assigned surveys, responses remained consistent over the 12-week study period (Figure 2). This implies that there is no drift within the sample and further indicates that our data collection was true to the original. Drift, in either a positive or negative direction, would indicate a loss of reliability in the VHI-10 instrument with collection of repeat measures. The absence of drift in these patients who were not undergoing therapy indicates the accuracy and stability of the VHI-10 over the course of time.

Of the collected surveys, 362 were collected in the home setting. These showed a mean VHI of 10. The number of surveys in the home setting that differed from the in-office score by at least the MCID was 231, which is 64% of the total number of collected scores. Of these, 54 surveys (23%) were higher than the office score by the MCID while 177 surveys (77%) were lower than the in-office score by the MCID ($p < .001$) (Figure 3).

In the work setting, 177 surveys were collected. These showed a mean VHI of 13.9. The number of surveys in the work setting that differed from the in-office score by at least the MCID was 64, which is 34% of the total number of collected scores. Of these, 20 surveys (31%) were higher than the in-office score by the MCID of 6 while 44 surveys (69%) were lower than the in-office score by the MCID ($p < .01$) (Figure 4).

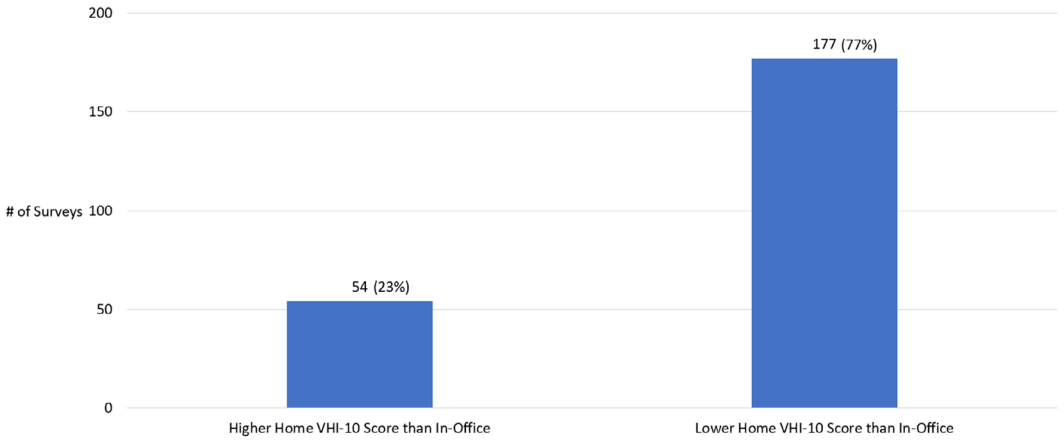

FIGURE 3 Home versus in-office survey results.

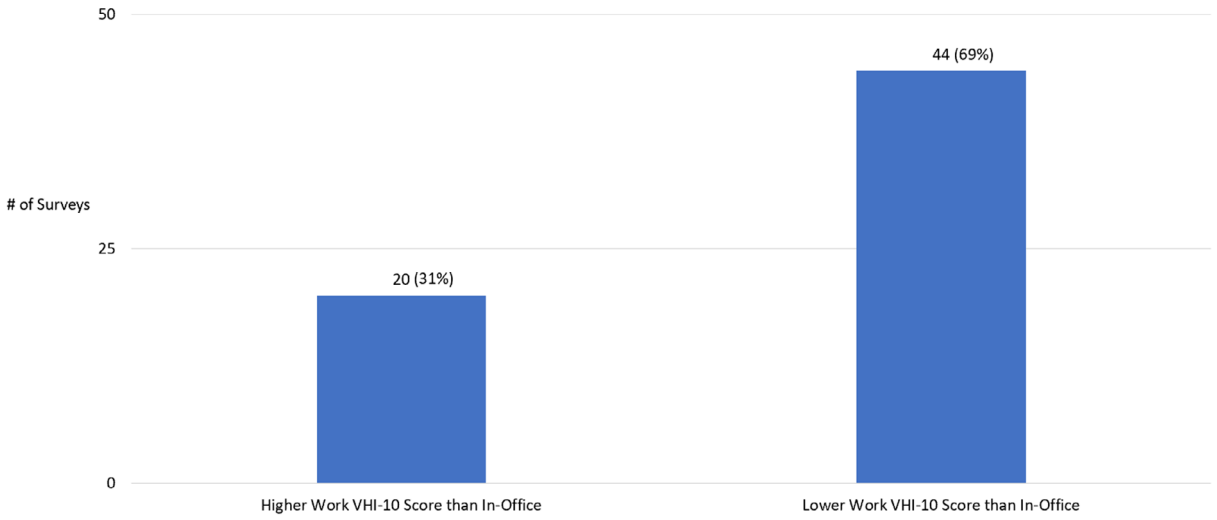

FIGURE 4 Work versus in-office survey results.

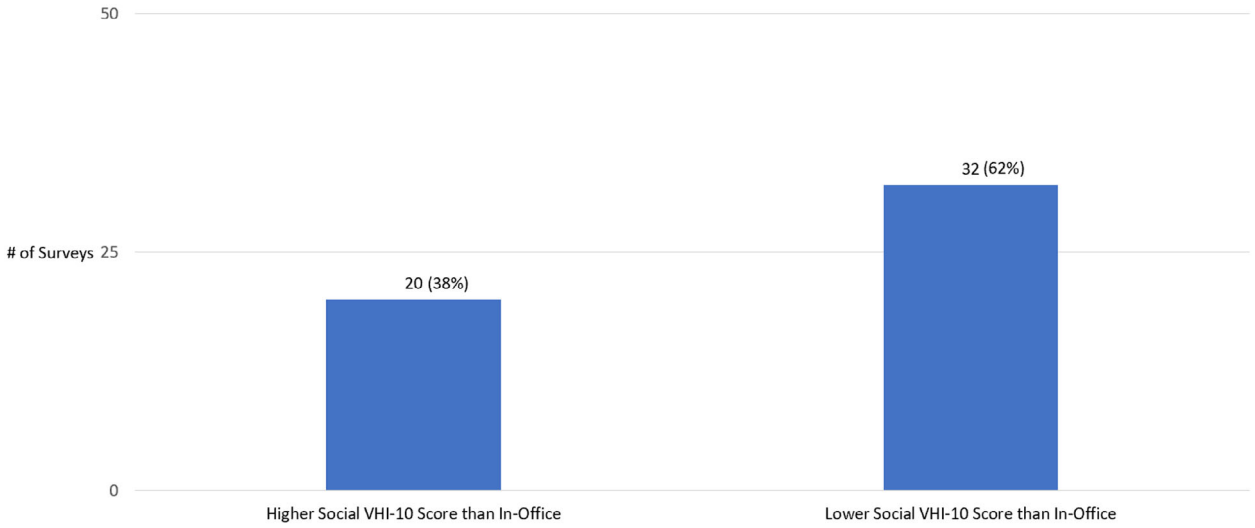

FIGURE 5 Social versus in-office survey results.

In the social setting, 97 surveys were collected. These showed a mean VHI of 13.8. The number of surveys in the social setting that differed from the in-office score by at least the MCID was 52, which is 54% of the total number of collected scores. Of these, 20 surveys (38%) were higher than the in-office score by the MCID of 6 while 32 surveys (62%) were lower than the in-office score by the MCID ($p = .09$) (Figure 5).

4 | DISCUSSION

The results of this study demonstrate that 63% of collected ambulatory VHI scores differed significantly, in terms of clinical response (MCID), from those collected in the in-office settings. Of these, across all ambulatory settings, the majority are actually lower than the in-office baseline score. When considering the in-office VHI as an outcome measure of an intervention, our findings imply that ambulatory VHI-10 surveys may frequently mis-represent the degree of dysphonia patients are experiencing on a daily basis.

There are a myriad of theories as to why a VHI-10 score might be different in the office versus in ambulatory settings. Chief among them is whether the patient is more focused on their dysphonia while in the otolaryngology office, as the patient is in the office for the purposes of evaluation. As such, the higher VHI-10 scores demonstrated in this sample may reflect a patient's focus on their dysphonia. In the ambulatory or social setting, patients may be focused on other things and provide a different view of their dysphonia. Given this limitation, and the setting of the VHI-10 validation studies, VHI-10 questionnaires completed in the office are likely to most accurately measure subjective dysphonia.

Future work should continue to explore the role of setting on patient reported outcome measures. Validation studies are commonly performed in one setting, that is, the physician's office, but are rarely replicated in other settings where the PROM may be nevertheless utilized. One way to conceptualize the clinical impact of current study designs is to consider the diagnosis of White Coat Hypertension. There is a strong evidence base demonstrating that factors such as location (home vs. physician's office), presence of medical students, and whether the practitioner is a doctor or a nurse have a significant impact on results of blood pressure measurements.[11,12] By delineating these factors, clinicians may interpret blood pressure findings within their context and understand the appropriate treatment options. This multi-setting validation process is generally lacking in the arena of PROMs.

In the COVID-19 era, we encourage our patients to complete forms prior to coming to the office and rely heavily upon telemedicine and remote care delivery methods. Clinicians need to understand whether a PROM completed at home should be interpreted in the same way as one completed in the office. If ambulatory PROM responses differ from in-office PROM responses, as implied by our study, the PROM may not provide the precise information clinicians are seeking. Interpretation of PROM results may, therefore, be inaccurate and lead to inappropriate treatment plans and outcomes monitoring.

There are limitations of this study that are important to note. To our knowledge, this is the first study to assess the impact of environment on VHI-10 results. Patients who had active treatment, whether surgical or voice therapy, during the study period were excluded. It is, therefore, possible that the results presented here would not be representative of a surgical or voice therapy cohort. Second, not all patients completed the full complement of 36 surveys, which may limit the generalizability of the results. Third, we did not independently verify the veracity of the locations in which patients reported filling out their surveys. It is, therefore, possible that patients may have mis-represented the setting in which a given survey was completed. Lastly, we did not control for the time of day or other factors when collecting ambulatory surveys. It is, therefore, possible that external circumstances impacted the results of the study and were not accounted for.

5 | CONCLUSIONS

The finding that 63% of ambulatory surveys differ by the minimal clinically important difference from in-office surveys implies that the setting in which the VHI-10 is completed is a relevant factor in clinical interpretation. This finding underscores the need for additional investigations of Patient Reported Outcome Measures to best understand their ideal clinical applications and utility in outcomes monitoring. Only the comparison of results collected from patients in the same setting is likely valid when determining the patient reported severity of dysphonia.

ACKNOWLEDGMENTS

The authors acknowledge Dr. Stanislaw Sobotka, PhD for his assistance with statistical analysis.

CONFLICT OF INTEREST

The authors have no financial interests or conflicts of interests to disclose.

ORCID

Eliezer C. Kinberg https://orcid.org/0000-0001-5954-0237
Amir Taree https://orcid.org/0000-0003-4960-9709
Mingyang Gray https://orcid.org/0000-0002-1146-9942
Peter Filip https://orcid.org/0000-0002-6898-6633
Mark Courey https://orcid.org/0000-0001-6885-2929

REFERENCES

1. Francis DO, Daniero JJ, Hovis KL, et al. Voice-related patient-reported outcome measures: a systematic review of instrument development and validation. *J Speech Lang Hear Res*. 2017;60(1):62-88. doi: 10.1044/2016_JSLHR-S-16-0022
2. Carroll TL, Lee SE, Lindsay R, Locandro D, Randolph GW, Shin JJ. Evidence-based medicine in otolaryngology, part 6: patient-reported

outcomes in clinical practice. *Otolaryngol—Head Neck Surg (U S)*. 2018;158(1):8-15. doi:10.1177/0194599817731759
3. Basch E, Deal AM, Kris MG, et al. Symptom monitoring with patient-reported outcomes during routine cancer treatment: a randomized controlled trial. *J Clin Oncol*. 2016;34(6):557-565. doi:10.1200/JCO.2015.63.0830
4. Franklin PD, Chenok KE, Lavalee D, et al. Framework to guide the collection and use of patient-reported outcome measures In the learning healthcare system. *eGEMs (Generating Evid Methods Improv Patient Outcomes)*. 2017;5(1):17. doi:10.5334/egems.227
5. Greenhalgh J, Gooding K, Gibbons E, et al. How do patient reported outcome measures (PROMs) support clinician-patient communication and patient care? A realist synthesis. *J Patient-Rep Outcomes*. 2018;2:2. doi:10.1186/s41687-018-0061-6
6. Rosen C, Lee A, Osborne J, Zullo T, Murry T. Development and validation of the voice handicap index-10. *Laryngoscope*. 2004;114(9):1549-1556. doi:10.1097/00005537-200409000-00009
7. Wong E, Smith M, Stone DB, Palme CE, Smith MC, Riffat F. Arytenoid vertical height discrepancy in predicting outcomes after unilateral vocal cord medialization. *Laryngoscope*. 2019;130:418-422. doi:10.1002/lary.27900
8. Misono S, Yueh B, Stockness AN, House ME, Marmor S. Minimal important difference in Voice Handicap Index-10. *JAMA Otolaryngol—Head Neck Surg*. 2017;143(11):1098-1103. doi:10.1001/jamaoto.2017.1621
9. Sedaghat AR. Understanding the minimal clinically important difference (MCID) of patient-reported outcome measures. *Otolaryngol—Head Neck Surg (U S)*. 2019;161(4):551-560. doi:10.1177/0194599819852604
10. Vickers AJ. Validation of patient-reported outcomes: a low bar. *J Clin Oncol*. 2019;37(23):1990-1992. doi:10.1200/JCO.19.01126
11. Clark CE, Horvath IA, Taylor RS, Research JLC. Doctors record higher blood pressures than nurses: systematic review and meta-analysis. *Br J Gen Pract*. 2014;64:e223-e232. doi:10.3399/bjgp14X677851
12. Matthys J. In general practice, doctors record higher blood pressures in the presence of students. *Br J Gen Pract*. 2014;64(623):277. doi:10.3399/bjgp14X680029

How to cite this article: Kinberg EC, Taree A, Gray M, Filip P, Courey M. Patient reported outcome measures: The impact of environment on VHI-10 responses. *Laryngoscope Investigative Otolaryngology*. 2023;8(1):150-155. doi:10.1002/lio2.1016