VALUE IN HEALTH ▮ (2017) ▮▮▮–▮▮▮

Available online at www.sciencedirect.com

ScienceDirect

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

Reliability, Validity, and Responsiveness of InFLUenza Patient-Reported Outcome (FLU-PRO©) Scores in Influenza-Positive Patients

John H. Powers III, MD[1,], Elizabeth D. Bacci, PhD[2], M. Lourdes Guerrero, MD, MS[3], Nancy Kline Leidy, PhD[4], Sonja Stringer, MPH[4], Katherine Kim, MPH[4], Matthew J. Memoli, MD, MS[5], Alison Han, MD, MS[5], Mary P. Fairchok, MD[6,7,8], Wei-Ju Chen, PhD[7,8], John C. Arnold, MD[9], Patrick J. Danaher, MD[10], Tahaniyat Lalani, MBBS[8,11], Michelande Ridoré, MS[12], Timothy H. Burgess, MD, MPH[7,13], Eugene V. Millar, PhD[7,8], Andrés Hernández, MD, MS[14], Patricia Rodríguez-Zulueta, MD[15], Mary C. Smolskis, RN, MA[4], Hilda Ortega-Gallegos, BS[3], Sarah Pett, MD, PhD[16,17], William Fischer, MD[18], Daniel Gillor, DrMed[19], Laura Moreno Macias, MD[20], Anna DuVal, MPH[21], Richard Rothman, MD, PhD[21], Andrea Dugas, MD, PhD[21], Guillermo M. Ruiz-Palacios, MD, FIDSA[3]*

[1]*Clinical Research Directorate/Clinical Monitoring Research Program, Leidos Biomedical Research, Inc., NCI Campus at Frederick, Frederick, MD, USA;* [2]*Evidera, Seattle, WA, USA;* [3]*Instituto Nacional de Ciencias Médicas y Nutrición Salvador Zubirán, Mexico City, Mexico;* [4]*Evidera, Bethesda, MD, USA;* [5]*National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD, USA;* [6]*Madigan Army Medical Center, Fort Lewis, WA, USA;* [7]*Infectious Disease Clinical Research Program, Department of Preventive Medicine and Biostatistics, F. Edward Hébert School of Medicine, Uniformed Services University of the Health Sciences, Bethesda, MD, USA;* [8]*Henry M. Jackson Foundation for the Advancement of Military Medicine, Bethesda, MD, USA;* [9]*Naval Medical Center, San Diego, CA, USA;* [10]*Defense Institute for Medical Operations, San Antonio, TX, USA;* [11]*Naval Medical Center, Portsmouth, VA, USA;* [12]*Children's National Medical Center, Washington, DC, USA;* [13]*Walter Reed National Military Medical Center, Bethesda, MD, USA;* [14]*Instituto Nacional de Enfermedades Respiratorias Ismael Cosío Villegas, Mexico City, Mexico;* [15]*Hospital General Dr. Manuel Gea González, Mexico City, Mexico;* [16]*University College London, London, UK;* [17]*Kirby Institute, University of New South Wales, Sydney, New South Wales, Australia;* [18]*University of North Carolina, Chapel Hill, NC, USA;* [19]*Cologne, Germany;* [20]*Hospital General de Agudos JM Ramos Mejia, Buenos Aires, Argentina;* [21]*Johns Hopkins University School of Medicine, Baltimore, MD, USA (for the INSIGHT Flu 002 Plus Writing Group, the México Emerging Infectious Diseases Clinical Research Network, and the Infectious Diseases Clinical Research Program)*

ABSTRACT

Objectives: To assess the reliability, validity, and responsiveness of InFLUenza Patient-Reported Outcome (FLU-PRO©) scores for quantifying the presence and severity of influenza symptoms. **Methods:** An observational prospective cohort study of adults (≥18 years) with influenza-like illness in the United States, the United Kingdom, Mexico, and South America was conducted. Participants completed the 37-item draft FLU-PRO daily for up to 14 days. Item-level and factor analyses were used to remove items and determine factor structure. Reliability of the final tool was estimated using Cronbach α and intraclass correlation coefficients (2-day reliability). Convergent and known-groups validity and responsiveness were assessed using global assessments of influenza severity and return to usual health. **Results:** Of the 536 patients enrolled, 221 influenza-positive subjects comprised the analytical sample. The mean age of the patients was 40.7 years, 60.2% were women, and 59.7% were white. The final 32-item measure has six factors/domains (nose, throat, eyes, chest/respiratory, gastrointestinal, and body/systemic), with a higher order factor representing symptom severity overall (comparative fit index = 0.92; root mean square error of approximation = 0.06). Cronbach α was high (total = 0.92; domain range = 0.71–0.87); test-retest reliability (intraclass correlation coefficient, day 1–day 2) was 0.83 for total scores and 0.57 to 0.79 for domains. Day 1 FLU-PRO domain and total scores were moderately to highly correlated (≥0.30) with Patient Global Rating of Flu Severity (except nose and throat). Consistent with known-groups validity, scores differentiated severity groups on the basis of global rating (total: $F = 57.2$, $P < 0.001$; domains: $F = 8.9–67.5$, $P < 0.001$). Subjects reporting return to usual health showed significantly greater ($P < 0.05$) FLU-PRO score

Conflicts of interest: The contents of this publication are the sole responsibility of the author(s) and do not necessarily reflect the views, opinions, or policies of the Department of Health and Human Services, Uniformed Services University of the Health Sciences, the Department of Defense, or the Departments of the Army, Navy, or Air Force. Mention of trade names, commercial products, or organizations does not imply endorsement by the US government.

* *Address correspondence to:* John H. Powers III, Leidos Biomedical Research, Inc., in support of Division of Clinical Research, National Institute of Allergy and Infectious Diseases, 5601 Fishers Lane, Room 4D50, Bethesda, MD 20892.

E-mail: john.powers@nih.gov.

http://dx.doi.org/10.1016/j.jval.2017.04.014

improvement by day 7 than did those who did not, suggesting score responsiveness. **Conclusions:** Results suggest that FLU-PRO scores are reliable, valid, and responsive to change in influenza-positive adults.

Keywords: influenza, patient-reported outcome, psychometric, reliability, responsiveness, validity.

Introduction

Approximately 5% to 20% of the US population is infected with influenza yearly, with 200,000 hospitalizations and 36,000 deaths [1–3]. Worldwide, influenza causes 3 million to 5 million severe cases and 250,000 to 500,000 deaths annually [4]. Symptoms range from mild to severe and include various systemic and respiratory symptoms, with gastrointestinal symptoms occurring less frequently [3].

Despite the prevalence of influenza and many research studies evaluating its natural history and treatment options, there are few validated patient-reported outcome measures for quantifying symptoms. Two previously developed instruments have been published but are limited by the populations studied with smaller numbers of patients with influenza studied compared with influenza-like illness (ILI) [5,6].

A validated, standardized patient-reported influenza symptom scale that comprehensively assesses the symptom experience in influenza across multiple body systems would allow for consistent, accurate assessments of symptoms associated with various viral strains over the course of the disease within and across subgroups. Use would facilitate meta-analyses, cross-product evaluations, and more precise estimates of treatment effects. Standardized measures should be developed using good research practices [7–9]. Instruments intended for use in drug development should address recommendations of the US Food and Drug Administration [10], including attention to content validity and quantitative testing in the target population for designated contexts of use.

The purpose of the InFLUenza Patient-Reported Outcome (FLU-PRO©) measure is to comprehensively assess the presence and severity of influenza symptoms across body systems often affected by these viruses. The ultimate intent was to develop a reliable, valid, and responsive measure for use in profiling the symptomatic manifestations of influenza on any given day, track changes over time, and test the effects of treatments. To ensure content validity, we used a two-stage qualitative instrument development methodology. In stage 1, we conducted concept elicitation interviews in the United States and Mexico to gather information regarding patient experience of influenza symptoms (i.e., type, magnitude, expression, pattern of onset, and recovery) [11,12]. Results informed the development of a draft measure, including content (candidate items), structure (response options, recall, and instructions), and conceptual framework [13]. In stage 2, we conducted cognitive interviews to assess completeness, comprehension, and ease of use of the draft measure from the respondent's perspective [13]. This work resulted in a draft instrument with 37 candidate questions ready for quantitative testing in the target population.

The objectives of this study were to 1) evaluate performance of the 37 candidate items; 2) reduce the number of items as empirically and conceptually appropriate; 3) finalize measurement/domain structure and develop a scoring algorithm for the final instrument, the FLU-PRO; and 4) explore the reliability, validity, and responsiveness of FLU-PRO total and domain scores.

Methods

Study Design and Sample

This was a prospective, observational study of English- and Spanish-speaking hospitalized and nonhospitalized adults 18 years or older with acute influenza. Patients seeking care for influenza symptoms at participating military or civilian clinics in the United States (16 sites), Argentina (2 sites), the United Kingdom (1 site), and Mexico (3 sites) were recruited in influenza seasons in northern and southern hemispheres. Influenza status was assessed through a positive polymerase chain reaction, rapid antigen test, and/or viral culture by nasal or nasopharyngeal swab.

We prespecified subjects testing positive for influenza as the target population and the primary analytical sample, with a goal of 200 or more subjects (100 for confirmatory factor analysis [CFA] and 185 [5 per item] for exploratory factor analysis [EFA]) [14], assuming that 50% of enrolled subjects testing positive for influenza would permit separate analyses on performance of the FLU-PRO in ILI. Given the different context of use, ILI results are presented elsewhere.

A total of 536 English- and Spanish-speaking patients were enrolled in the study; 441 had diary entries on day 1 and at least 1 day thereafter, qualifying them for analyses. Two hundred twenty-one were influenza-positive (see Appendix Figure S2 in Supplemental Materials found at http://dx.doi.org/10.1016/j.jval.2017.04.014).

Procedures

Clinical research coordinators recruited participants with influenza-like symptoms. Patients providing consent: 1) completed clinic-based baseline assessments of sociodemographic and clinical characteristics; 2) were tested for laboratory-confirmed influenza; and 3) completed daily diaries for up to 14 days after enrollment. This included the 37-item draft FLU-PRO symptom diary and nine additional questions for validation purposes. At Mexico sites, diaries were completed via telephone interviews with data entered directly into a Web-based portal. Participants in the United States, the United Kingdom, and Argentina completed the survey via either an interviewer-administered method or a Web-based system using their personal devices. Translation procedures for Spanish followed the International Society for Pharmacoeconomics and Outcomes Research guidelines [15]. The study was conducted with informed consent, institutional review board approval, and in accordance with the Declaration of Helsinki [15].

Instruments: Patient-Reported Outcomes

InFLUenza Patient-Reported Outcome

The draft FLU-PRO Questionnaire instructed respondents to rate the severity of 37 influenza symptoms over the past 24 hours, including those related to the nose, throat, eye, chest, head, stomach, fatigue, and body aches/pains. Six items measured the same symptom using different wording to select the best performing item for the final instrument. For 32 of the 37 items, respondents rated the severity of each symptom on five-point Likert-type scales, with 0 indicating "Not at all"; 1, "A little bit"; 2, "Somewhat"; 3, "Quite a bit"; and 4, "Very much." For the five remaining items, severity is expressed as frequency of occurrence: vomiting or diarrhea (0 time, 1 time, 2 times, 3 times, or 4 or more times). Sneezing, coughing, and coughed-up mucus or phlegm were expressed on a scale from 0 ("Never") to