

Research Article

Concordance between Patients and Clinicians for Reporting Symptoms Associated with Treatment for Chronic Hepatitis C during a Pragmatic Clinical Trial

Bryce B. Reeve¹, Larry Michael,² Joy A. Peter,³ Paul W. Stewart,⁴ Anna S.-F. Lok,⁵ Meichen Dong,⁴ Ken Bergquist,² David R. Nelson,³ and Donna M. Evon⁶

¹Center for Health Measurement, Department of Population Health Sciences, Duke University School of Medicine, Durham, NC, USA

²Center for Gastrointestinal Biology and Disease, School of Medicine, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

³Division of Gastroenterology, Hepatology & Nutrition, Department of Medicine, University of Florida, Gainesville, FL, USA

⁴Department of Biostatistics, Gillings School of Global Public Health, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

⁵Division of Gastroenterology and Hepatology, Department of Internal Medicine, University of Michigan, Ann Arbor, MI, USA

⁶Division of Gastroenterology and Hepatology, School of Medicine, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

Correspondence should be addressed to Bryce B. Reeve; bryce.reeve@duke.edu

Received 30 December 2021; Accepted 20 June 2022; Published 4 July 2022

Academic Editor: Soe Thiha Maung

Copyright © 2022 Bryce B. Reeve et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Background. Despite high efficacy rates for direct acting antiviral regimens to cure hepatitis C virus infection, many patients experience treatment-related symptoms. Accurate reporting of adverse events is mandatory to determine drug safety. Previous research in other medical conditions has documented discordance between clinician-reported and patient-reported symptomatic adverse events. **Aims.** To explore concordance and associated factors, between clinician-recorded and patient-reported fatigue, headache, and nausea/vomiting during a clinical trial of three treatment regimens. **Methods.** Data were collected between treatment start and 31 days posttreatment. Patients completed Patient-Reported Outcomes Measurement Information System measures of fatigue and nausea/vomiting and the Headache Impact Test. Clinician-recorded data were abstracted from medical records. Concordance was evaluated by weighted kappa. Demographic and clinical factors associated with concordance were identified using logistic regression models. **Results.** Participants included 1,058 patients treated for chronic hepatitis C (average 54.9 years; 43% Black; 59% male). Weighted kappa estimates and 95% confidence intervals between patients (no/mild vs. moderate/severe symptoms) and clinicians (not present vs. present) were fatigue ($k = 0.09$, 0.02-0.16), headache ($k = 0.08$, 0.02-0.14), and nausea/vomiting ($k = 0.20$, 0.11-0.28). Older age and having private insurance (compared to Medicaid) were associated with better headache concordance. Older age, male, absence of psychiatric condition, and ≤ 2 comorbidities were associated with better nausea/vomiting concordance. **Conclusions.** Poor concordance was observed between patient-reported and clinician-recorded symptomatic adverse events. Despite study limitations, previous literature in other conditions support these findings. Integrating patient-reported data to inform adverse event reporting would improve evaluations of treatment safety (<http://CT.gov/> Registration: NCT02786537).

1. Introduction

Chronic hepatitis C virus (HCV) infection is a potentially life-threatening disease that affects more than 2.4 million people living in the United States (US) [1]. Left untreated, HCV can lead to cirrhosis, liver failure, liver cancer, and death [2]. Many patients living with chronic HCV experience extrahepatic symptoms including several somatic, gastrointestinal, and neuropsychiatric symptoms that impair quality of life compared to patients with other liver diseases and the general US population [3–5].

Several all-oral, direct-acting antiviral (DAA) regimens have been approved since 2014 and are highly efficacious. Approximately 95% of patients treated will achieve viral cure after 8 to 24 weeks of DAA treatment [6]. The safety profiles show these new DAA regimens to be much better tolerated than the interferon-based regimens [7–9]. This study focuses on three symptoms (fatigue, nausea/vomiting, and headaches) that were most commonly reported as adverse events (AEs) during DAA treatment trials [10–12].

Phase III registration trials are designed to assess treatment benefit and safety. While many safety signals are identified by laboratory tests, symptomatic AEs are typically solicited and reported by clinical investigators. However, there is a growing literature on the value of integrating direct patient reporting to identify and monitor symptoms associated with medical treatments in clinical trials and routine clinical care. A number of studies have found that clinicians tend to underreport the frequency and severity of symptoms compared to patient self-report [13–18]. These findings have resulted in the development of symptom AE tracking systems to integrate the patients' voices in the assessment of treatment toxicity in clinical trials [19, 20]. For cancer therapies, the FDA has created a website to provide public information on the tolerability associated with cancer treatment using data collected directly from patients participating in oncology clinical trials [21].

The goal of this study was to explore concordance between clinician-recorded and patient-reported AEs concerning fatigue, headache, and nausea/vomiting during a pragmatic clinical trial of three DAA regimens for the treatment of chronic hepatitis C, genotype 1. Based on the literature in other chronic diseases [13–18], we expected to find poor-to-fair levels of agreement between clinicians and patients in part due to underreporting by clinicians relative to the patients' reports of these symptoms.

2. Methods

2.1. Study Design and Participants. The PRIORITIZE study was a multicenter, randomized, pragmatic clinical trial (<http://CT.gov/> Registration: NCT02786537) designed to compare three DAA regimens that were commonly used in clinical practices for treatment of HCV genotype 1, at the time this trial was launched in 2016 (1) ledipasvir/sofosbuvir (LDV/SOF), Harvoni®, Gilead Sciences, Foster City, CA; (2) elbasvir/grazoprevir (EBR/GZR), Zepatier™, Merck and Co, Whitehouse Station, NJ; and (3) paritaprevir/ritonavir/

ombitasvir + dasabuvir (PrOD), Viekira Pak/Viekira XR™, AbbVie Pharmaceuticals, Abbott Park, IL [22].

The PRIORITIZE trial design, efficacy rates and AEs based on laboratory tests and symptoms reported by clinicians have been previously published [22]. In brief, participants with chronic HCV genotype 1 were enrolled from 34 community and academic liver centers in the US which are part of the HCV-TARGET Network [23]. Participants were initially randomized 1:1:1 to LDV/SOF, EBR/GZR, or PrOD, stratified by cirrhosis status and genotype 1 subtype (a or b). After start of the trial, PrOD was discontinued, and all remaining participants were randomized to LDV/SOF or EBR/GZR. The enrollment target was 1600 participants. Adult participants (>18 yrs) with HCV infection (genotype 1a or 1b) who presented for antiviral treatment were invited to participate if, in their clinician's opinion, therapy with any of the study regimens was appropriate. Individuals were excluded for inability to provide written informed consent, current or historical evidence of hepatic decompensation (variceal bleeding, hepatic encephalopathy, or ascites) unless this was prior to successful liver transplant, Child-Turcotte-Pugh (CTP) stage B or C cirrhosis, pregnancy, or breastfeeding status, or health insurance drug formulary that did not approve of prescription for LDV/SOF which was unable to be provided by the study.

The frequency of follow-up visits and methods used to solicit and record AEs were based on standard of care for HCV therapy at each participating site. Treatment duration was typically 8 to 24 weeks depending on DAA regimen, patient clinical characteristics such as cirrhosis status, HCV genotype, and prior treatment failure, and insurance payer approval of prescribed duration. The study protocol was approved by the Institutional Review Board of all participating centers, and all patients provided written informed consent prior to enrollment.

2.2. Measures. Clinician-recorded and patient-reported symptom AEs were collected for the purpose of comparing the three treatment regimens on safety as a secondary outcome. Enrollment for the first patient began in June 2016, and the last data point was collected in August 2020. PRO surveys were administered through the use of technology/devices outside of clinical interactions using email prompted, web-based surveys via Research Electronic Data Capture System (REDCap) or via phone-based surveys conducted by trained staff at the University of Florida's Survey Center.

Patient-reported fatigue was assessed via the NIH's Patient-Reported Outcomes Measurement Information System® (PROMIS®) Fatigue short form 8a [24–27]. Surveys were included as completed by the patient without adjustment for missing values. The resulting PROMIS *T*-scores for fatigue were categorized as normal (<55), mild (55 to <60), moderate (60 to <70), and severe (70+).

Similarly, nausea/vomiting was evaluated via the PROMIS Nausea/Vomiting 4-item short form [28]. Surveys were included as completed by the patient without adjustment for missing values. The resulting PROMIS *T*-scores were categorized as none (<42), mild (42 to <60), moderate (60 to <70), and severe (70+).

Headache was assessed by the 6-item Headache Impact Test (HIT-6) [29, 30]. The HIT-6 score was computed as the sum of the 6 items. If more than 1 item had a missing value, then the HIT-6 score was treated as a missing value. If only one item had a missing value, then the HIT-6 score was imputed as 6 times the average of the available 5 items. The HIT-6 scores were categorized as no to little impact on life (<50), some impact (50-55), substantial impact (56-59), and very severe impact on life (>59).

The psychometric properties of these three PRO measures have been previously evaluated in a sample of patients with HCV undergoing DAA therapy [31].

For purposes of this analysis, we limited the window of PRO measures and medical-records data to dates between treatment start and end-of-treatment +31 days. Study participants may have completed PRO surveys on 1, 2, or 3 occasions during this treatment window. The PROMIS measures of fatigue and nausea/vomiting were completed once and twice by approximately 88% and 12% of the sample, respectively. On a different assessment schedule, the HIT-6 headache measure was completed once (17%), twice (74%), or three times (9%). For each PRO measure, we selected the participant's highest (worst) score for comparison to their clinicians' reporting of related symptomatic AEs.

The clinician-recorded data were abstracted from the patients' medical records. The schedule of clinic visits during DAA treatment was not specifically defined due to the nature of the pragmatic trial design. Clinical visits were based on standard of care for HCV therapy in patients with HCV genotype 1, with and without cirrhosis at the time the study was designed, with practitioners cognizant of adhering to the AASLD/IDSA guidelines for the treatment of chronic HCV in adult patients [6].

Extraction of clinician-recorded AEs from medical records were standardized using the Medical Dictionary for Regulatory Activities (MedDRA) and subsequently coded as either present or not present in the medical record. Where absent from the patient's medical record, AEs were assumed to have been deemed absent by the clinician. For fatigue, we coded "present" when the patient's medical record captured an event of fatigue, anemia, lethargy, or malaise. For nausea/vomiting, we coded "present" when the patient's medical record captured an event of nausea, vomiting, or hematemesis. For headache, we coded "present" when the patient's medical record captured an event of headache, migraine, or sinus headache. Selected AE terms were chosen by clinical colleagues based on similarity to patient-reported symptom or conditions where the symptom is a cardinal indicator (e.g., anemia and fatigue).

2.3. Data Analyses. The analyses were limited to estimation of concordance and hypothesis generation due to study design limitations, particularly the unknown temporal coincidence of the AE and PRO records. Thus, exploratory statistical methods were used to investigate concordance between clinician-recorded and patient-reported AEs. Hypothesis-testing methods and *p* values were *not* used. The statistical computations were performed using SAS software version 9.4 (SAS Institute Inc., Cary, NC, USA).

2.3.1. Descriptive Summaries. Qualitative and graphical statistical methods were used to visualize bivariate relationships and empirical frequency distributions of the outcome variables and to promote insights into the appropriateness of potential quantitative statistical approaches. Quantitative descriptive statistical methods were used to characterize the distributions of baseline characteristics, treatment assignments, the PROMIS *T*-scores, and HIT-6 scores relative to severity categories.

2.3.2. Evaluation of Agreement. Agreement between patient-reported symptom AE presence and clinician-report of the symptomatic AE was evaluated by estimation of weighted kappa (κ) coefficients. For that purpose, we dichotomized the severity of the patients' reports as "normal" versus "mild," "moderate," or "severe" to compare with the presence of a provider-recorded AE during the time a patient was on HCV therapy. We interpreted agreement as follows: poor ($\kappa < 0.20$), fair ($0.20 < \kappa < 0.40$), moderate ($0.40 < \kappa < 0.60$), good ($0.60 < \kappa < 0.80$), and very good ($\kappa > 0.80$) [32].

2.3.3. Predictive Models for Agreement. Logistic regression models were explored in which the dependent variable was an indicator of the occurrence of concordance between patient report and clinician recording for each symptom AE (fatigue, nausea/vomiting, and headaches). Logistic regression models that included one predictor at a time were used to estimate unadjusted odds ratios for age, sex, ethnicity, race, cirrhosis, BMI, insurance type, treatment regimen, ribavirin, and frequency of medical and psychiatric comorbidities.

These same predictors were included in a multivariable logistic regression model fitted via LASSO (least absolute shrinkage and selection operator) estimation in which the variable selection and regularization was based on the Schwarz-Bayesian Criterion [33]. This approach is designed to identify parsimonious, best-fitting models, instead of attempting to identify highly predictive variables. LASSO estimation is designed to cope with multicollinearity among predictors [34].

As a sensitivity analysis, a logistic regression model was used to examine if the level of concordance was associated with the time interval between the clinician report of a symptom AE and the patient self-report of the worst symptom experience.

3. Results

3.1. Sample Characteristics. Table 1 summarizes the demographic and clinical characteristics of the 1058 HCV patients included in this analysis. Age at enrollment ranged from 18 to 86 years (average 54.9 years), and 43% of patients were Black, and 59% were male. All patients had at least one medical comorbidity, 17% had cirrhosis, and 35% had one or more psychiatric conditions. A majority (58%) of patients received EBR/GZR, 14% of patients received ribavirin in addition to DAA, and 96% of those with available virologic outcome data had a sustained virologic response to

TABLE 1: Sample characteristics of 1058 adults with chronic HCV.

Characteristic	Statistic
Age	
Years—mean (SD)	54.9 (11.8)
Sex— <i>n</i> (%)	
Female	430 (40.6%)
Male	628 (59.4%)
Race— <i>n</i> (%)	
White	538 (50.9%)
Black	454 (42.9%)
Other	45 (4.2%)
Ethnicity— <i>n</i> (%)	
Hispanic	59 (5.6%)
Non-Hispanic	971 (91.8%)
Health insurance— <i>n</i> (%)	
Medicaid	419 (39.6%)
Medicare	196 (18.5%)
Private	375 (35.4%)
Uninsured	51 (4.8%)
Other	17 (1.6%)
Body mass index	
Mean (SD)	28.5 (6.3)
Obese category (≥ 30)— <i>n</i> (%)	386 (36.5%)
Medical comorbidities	
Mean (SD)	3.5 (1.7)
1-2 comorbidities— <i>n</i> (%)	368 (34.8%)
3-4 comorbidities— <i>n</i> (%)	411 (38.8%)
5+ comorbidities— <i>n</i> (%)	279 (26.4%)
Cirrhosis— <i>n</i> (%)	
No	881 (83.3%)
Yes	177 (16.7%)
Psychiatric comorbidities— <i>n</i> (%)	
None	686 (64.8%)
1+ comorbidities	372 (35.2%)
DAA treatment regimen— <i>n</i> (%)	
EBR/GZR	612 (57.8%)
LDV/SOF	316 (29.9%)
PrOD	130 (12.3%)
Ribavirin use— <i>n</i> (%)	
No	910 (86.0%)
Yes	148 (14.0%)
Virologic outcome— <i>n</i> (%) of those with virologic outcome	
Sustained virologic response	875 (96.2%)
Virologic failure	35 (3.8%)

Note: some percentages may not add to 100% as there was missing. Abbreviations: SD: standard deviation; DAA: direct acting antiviral; EBR/GZR: elbasvir/grazoprevir (Zepatier™); LDV/SOF: ledipasvir/sofosbuvir (Harvoni®); PrOD: paritaprevir/ritonavir/ombitasvir + dasabuvir (Viekira Pak/Viekira XR™).

treatment. Figure 1 provides a Consort diagram of how many patients consented to the study and completed PRO measures for fatigue, nausea/vomiting, and headaches.

3.2. Distribution of Patient-Reported Symptom Scores and Clinician Indication of Symptom Presence. Among 1002 participants who completed at least one PROMIS Fatigue measure, the average maximum fatigue score during treatment was 51.1 (SD = 10.7) with 63.2% of the participants with normal fatigue levels, 17.5% of the participants with mild fatigue, 15.9% of the participants with moderate fatigue, and 3.5% of the participants self-reporting severe fatigue levels. Among 1057 participants who completed at least one HIT-6 measure, the mean headache scores was 49.0 (SD = 10.6) during treatment with 53.9% of the participants self-reporting levels of headache consistent with no to little impact on life, 16.3% of the participants with some impact, 9.5% of the participants with substantial impact, and 20.3% of the participants self-reporting levels of headache associated with very severe impact on life. Among 970 participants who completed at least one PROMIS nausea/vomiting measure, the mean nausea/vomiting score during treatment was 45.6 (SD = 9.8) with 54.4% of the participants self-reporting no nausea/vomiting, 34.1% of the participants with mild nausea/vomiting, 9.9% with moderate nausea/vomiting, and 1.5% with severe nausea/vomiting levels.

Clinicians' reporting of symptom AE presence in the medical record for the corresponding patient-reported samples were 21.4% with fatigue, 15.6% with headache, and 12.6% with nausea or vomiting. For fatigue, use of ribavirin, insurance type, and treatment regimen were selected by the LASSO procedure as predictors for occurrence of a clinician-recorded symptom AE for fatigue. Patients receiving ribavirin in their DAA regimen were more likely to have a clinician-recorded fatigue event compared to patients who were not on ribavirin (OR = 5.47, 95%CI = 3.36-8.91). Patients receiving Medicaid insurance (OR = 1.51, 95%CI = 1.04-2.20) and patients with other insurance types (OR = 3.32, 95%CI = 1.13-9.80) were more likely to have a clinician-recorded fatigue event compared to those with private insurance. While selected by the LASSO procedure, the subsequent logistic regression models did not suggest differences among treatment regimens for predicting clinician-recorded fatigue events. For nausea/vomiting and for headache, no predictors were selected by the LASSO procedure.

3.3. Association between Clinician-Recorded and Patient-Reported Symptom AEs. Figure 2 presents the percentages for clinician recording of fatigue, headache, or nausea/vomiting in the medical record by patient-reported symptom severity level. Using fatigue as an example, 633 patients self-reported normal levels of fatigue and for 17% of these patients, the clinicians recorded fatigue in the medical record. For the 175 patients who self-reported mild levels of fatigue, 27% had a clinician-recorded fatigue event in their medical record. For the 159 patients who self-reported moderate fatigue levels, 29% of the patients had a clinician-recorded fatigue event. For the 35 patients who

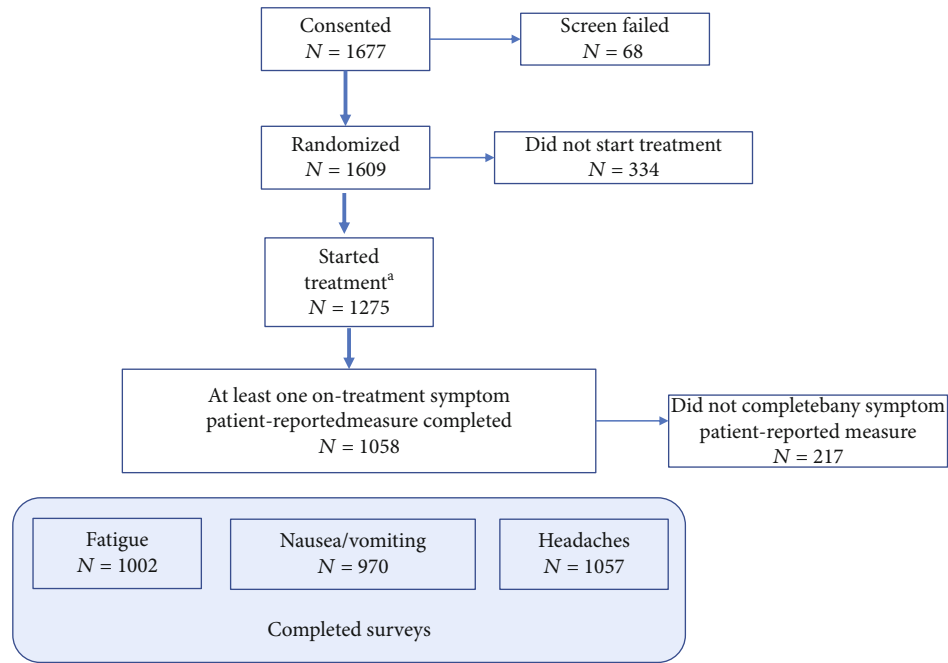


FIGURE 1: Consort diagram of how many patients consented to the study and completed patient-reported measures for fatigue, nausea/vomiting, and headaches.

self-reported severe levels of fatigue, only 31% had an event of fatigue recorded in their medical record.

To evaluate agreement in terms of weighted kappa coefficients, we dichotomized the patient-reported data in two ways. First, we set the threshold to differentiate no/normal symptom experience from the combination of mild, moderate, or severe symptom experiences. The weighted kappa estimates and 95% confidence intervals were fatigue ($k = 0.12$, 0.06-0.18), headache ($k = 0.07$, 0.02-0.12), and nausea/vomiting ($k = 0.17$, 0.12-0.22). Next, we set the threshold to differentiate no and mild symptom experiences combined from moderate and severe symptom experiences combined. The kappa coefficients estimates were fatigue ($k = 0.09$, 0.02-0.16), headache ($k = 0.08$, 0.02-0.14), and nausea/vomiting ($k = 0.20$, 0.11-0.28).

3.4. Prediction of Concordance between Patient-Reporting and Clinician-Recording of Symptom AEs. Table 2 summarizes the odds ratios for demographic and clinical factors selected by the LASSO procedure as predictors of concordance in reporting of symptom AEs between patients and clinicians. No predictors were selected for fatigue. For headache, older age was associated with better concordance, and having Medicaid insurance (compared to private insurance) was associated with worse concordance. For nausea/vomiting, older age was associated with better concordance, whereas being female (compared to male) having a psychiatric condition (versus not) and having 3 to 4 or 5+ comorbidities (compared to ≤ 2 comorbidities) was associated with worse concordance. Both insurance type and treatment regimen were selected by the LASSO procedure but the odds ratios in the subsequent logistic regression model were not significant for nausea/vomiting. In exploratory/sensitivity

analyses, there was no evidence for an association between the level of concordance and the time interval between the patient self-report and the clinician report (fatigue; $p = 0.13$; headache, $p = 0.74$; nausea/vomiting, $p = 0.31$).

4. Discussion

In our exploratory analyses, we investigated concordance between patient-reported and clinician-recorded symptom AEs in a cohort of patients participating in a pragmatic trial of DAA treatments for chronic HCV infection. Even when setting the threshold for patient-report of symptom experiences at moderate or severe, the level of agreement between patients and clinicians was categorized as *poor* for headache (weighted kappa = 0.08), fatigue (kappa = 0.09), and nausea/vomiting (kappa = 0.20). Consistently, clinicians reporting of any of these three symptoms in the medical record were much less frequent than what patients reported. For example, 194 patients reported moderate to severe fatigue; yet only 29% of the time the clinicians recorded a fatigue event in the medical record for these patients. For headache, 315 patients reported moderate to severe levels yet, only 21% of their medical records included a headache AE recorded by the clinicians. For nausea/vomiting, 111 HCV patients indicated moderate to severe symptom levels; yet only 31% of their records included the corresponding AE notation by the clinician.

These findings of low agreement and lower percentages of clinicians' recognition of symptoms are consistent with findings in other health conditions or medical treatments. For example, Xiao et al. (2013) published a review of 36 studies looking at agreement between patients and doctors for patients undergoing cancer treatment. Consistently

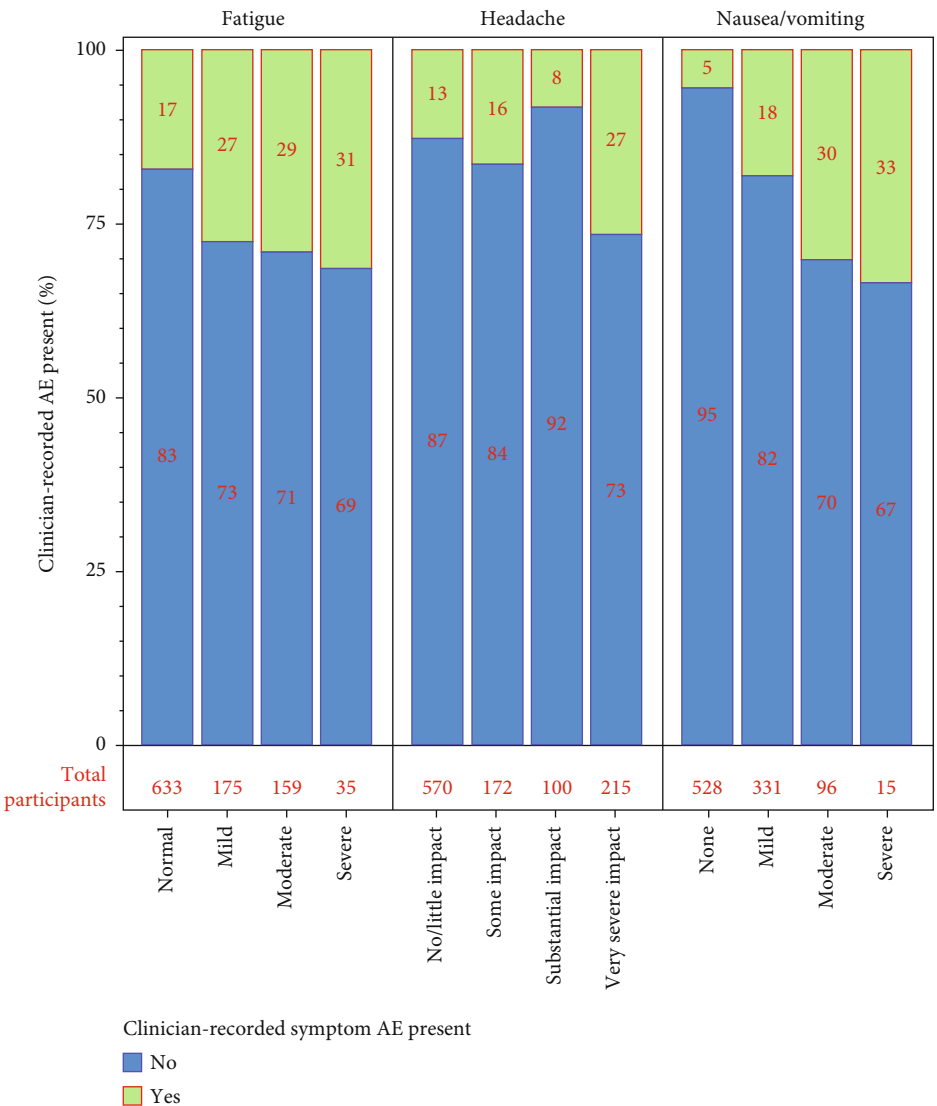


FIGURE 2: Clinician recording of fatigue, headache, and nausea/vomiting in the medical record by patient-reported symptom severity level.

across studies, Xiao et al. (2013) found that clinicians underestimate the presence, severity, and impact of symptoms compared to patients' self-report, even when symptoms were severe and distressing to patients [16]. McColl et al. examined agreement data from four randomized clinical trials (RCTs) including 2674 patients receiving treatment for gastroesophageal reflux disease. Agreement levels after treatment were in the fair to good range (kappas ranged from 0.31 to 0.73), with poorer agreement for heartburn and epigastric pain and better agreement for dysphagia [35]. Barbara et al. examined agreement in 176 participants in a RCT evaluating vaccines for influenza. Agreement for symptoms ranged from kappas of 0.05 (chills) to 0.51 (earache). Agreement for fatigue was poor ($\kappa=0.13$) and for headache was fair ($\kappa=0.21$) with clinicians underreporting the symptoms [36]. Justice et al. examined agreement data from a RCT of antiviral therapy among 1262 patients with moderate HIV disease. While agreement was higher in this study (kappas ranged from 0.50 to 0.80), clini-

cians "substantially under reported the prevalence of symptoms when compared to patients." (p. 401) [37] For example, fatigue was reported by 38.7% of clinicians and 64.7% of patients. Headache was reported by 25.3% of clinicians and 53.0% of patients. Nausea was reported by 12.1% of clinicians and 32.8% of patients.

For headache and nausea/vomiting, our exploratory analyses suggested that older age predicted better concordance. This may reflect either better communication between patients and doctors for older patients or that the clinicians may expect higher symptom burden for their older patients [38]. More research is needed in this area. For nausea/vomiting, a history of psychiatric disorders predicted less concordance relative to not having a history of psychiatric disorders. While this finding needs to be studied further, it may be because patients with psychiatric disorders are less willing to communicate their symptoms or that clinicians may unconsciously disregard the symptom complaints of patients with mental

TABLE 2: Prediction of concordance between patient-reported and clinician-recorded adverse events via logistic regression models selected by LASSO estimation.

	Fatigue Odds ratio (95% CI)	Headache Odds ratio (95% CI)	Nausea/vomiting Odds ratio (95% CI)
Age			
In 10-year units	—	1.36 (1.21, 1.52)	1.37 (1.18, 1.55)
Sex			
Female vs. male	—	—	0.59 (0.45, 0.77)
Psychiatric comorbidities			
Yes vs. no	—	—	0.64 (0.48, 0.85)
Medical comorbidities			
3 or 4 vs. ≤ 2	—	—	0.58 (0.41, 0.82)
≥ 5 vs. ≤ 2	—	—	0.61 (0.41, 0.90)
Insurance type			
Uninsured vs. private	—	1.32 (0.71, 2.46)	0.83 (0.44, 1.59)
Other vs. private	—	0.95 (0.35, 2.57)	0.64 (0.23, 1.81)
Medicare vs. private	—	0.82 (0.57, 1.20)	0.98 (0.65, 1.49)
Medicaid vs. private	—	0.75 (0.56, 0.99)	0.81 (0.59, 1.11)
Treatment regimen			
LDV/SOF vs. EBR/GZR	—	—	1.18 (0.86, 1.62)
PrOD vs. EBR/GZR	—	—	1.07 (0.71, 1.62)

Note: each of the three dichotomous outcome variables indicates presence/absence of concordance between patient and clinician reports. Predictor variables examined included patient age, sex, race, ethnicity, BMI, insurance type, medical comorbidities, cirrhosis status, psychiatric comorbidities, treatment regimen, ribavirin use, and SVR status. Empty cells represent predictor variables that were not selected for inclusion in the LASSO logistic regression model. Abbreviations: CI: confidence interval; LASSO: least absolute shrinkage and selection operator; PrOD: paritaprevir/ritonavir/ombitasvir + dasabuvir (Viekira Pak/Viekira XR™); EBR/GZR: elbasvir/grazoprevir (Zepatier™); LDV/SOF: ledipasvir/sofosbuvir (Harvoni®).

illness compared to patients without known psychiatric disorders. Similarly, the role of sex, medical comorbidities, health insurance, and treatment regimen needs to be studied further.

A limitation of this exploratory study is that surveys of patients' symptoms were conducted independently of clinical encounters. Our exploratory evaluation did not find an association between reporting periods between the patients' report and clinicians' report with level of concordance. In addition, we cannot confirm that patients mentioned these symptoms to their clinicians. We hope future studies in the hepatology field will explore these associations more thoroughly and include a broader range of AE symptoms during clinical trials. However, even with better study designs, we anticipate findings would be consistent with findings from multiple other studies reporting underreporting of AEs by clinicians [13–18, 35–37]. Another limitation is that this study was limited to English-speaking participants in the United States; thus, it is unknown if these results will generalize to non-English speaking patients and patients in other nations with different healthcare systems and different cultures.

4.1. Conclusions. This is the first study we are aware of to explore the concordance between patient-reported symptoms using surveys and clinician recording of the corresponding events captured in medical records in a population of patients with hepatitis C undergoing DAA

treatment. The findings have implications not only for future trials of hepatitis C treatment but will be highly relevant to future drug trials and subsequent comparative effectiveness studies for all liver disease populations. This study suggests a poor level of concordance between patient-reported and clinician-recorded symptom AEs with clinicians tending to underreport symptom AEs. This finding is consistent with the literature in many other health conditions and medical treatments [13–18, 35–37] and supports the increasing need to evaluate symptoms directly using patient-reported measures, to obtain a more accurate depiction of safety and symptoms during treatment with investigational as well as approved treatments.

Data Availability

Data comes from a registered clinical trial (<http://CT.gov/Registration:NCT02786537>).

Ethical Approval

The authors confirm that the ethical policies of the journal, as noted on the journal's author guidelines page, have been adhered to, and the appropriate ethical review committee approval has been received. The study conformed to the US Federal Policy for the Protection of Human Subjects.

Conflicts of Interest

Bryce Reeve served as a consultant for the University of Florida. Anna Lok receives research funding from Gilead Sciences, Inc.

Acknowledgments

This study was funded in part by the Patient-Centered Outcomes Research Institute (PCORI) Award (HPC-1503-27891).

References

- [1] S. Blach, S. Zeuzem, M. Manns et al., "Global prevalence and genotype distribution of hepatitis C virus infection in 2015: a modelling study," *The Lancet Gastroenterology & Hepatology*, vol. 2, no. 3, pp. 161–176, 2017.
- [2] Centers for Disease Control and Prevention, "Hepatitis C questions and answers for health professionals," <https://www.cdc.gov/hepatitis/hcv/hcvfaq.htm>.
- [3] S. Monaco, S. Ferrari, A. Gajofatto, G. Zanusso, and S. Mariotto, "HCV-related nervous system disorders," *Clinical and Developmental Immunology*, vol. 2012, Article ID 236148, 9 pages, 2012.
- [4] G. R. Foster, R. D. Goldin, and H. C. Thomas, "Chronic hepatitis C virus infection causes a significant reduction in quality of life in the absence of cirrhosis," *Hepatology*, vol. 27, no. 1, pp. 209–212, 1998.
- [5] Z. Younossi, J. Kallman, and J. Kincaid, "The effects of HCV infection and management on health-related quality of life," *Hepatology*, vol. 45, no. 3, pp. 806–816, 2007.
- [6] American Association for the Study of Liver Diseases, "AASLD-DISA recommendations for testing, managing, and treating hepatitis C," <http://www.hcvguidelines.org/>.
- [7] E. Lawitz, A. Mangia, D. Wyles et al., "Sofosbuvir for previously untreated chronic hepatitis C infection," *New England Journal of Medicine*, vol. 368, no. 20, pp. 1878–1887, 2013.
- [8] A. J. Muir, "The rapid evolution of treatment strategies for hepatitis C," *Official Journal of the American College of Gastroenterology ACG*, vol. 109, no. 5, pp. 628–635, 2014.
- [9] J. Sperl, G. Horvath, W. Halota et al., "Efficacy and safety of elbasvir/grazoprevir and sofosbuvir/pegylated interferon/ribavirin: a phase III randomized controlled trial," *Journal of Hepatology*, vol. 65, no. 6, pp. 1112–1119, 2016.
- [10] Gilead Sciences, Inc, "HARVONI drug label," https://www.gilead.com/~media/files/pdfs/medicines/liver-disease/harvoni/harvoni_pi.pdf.
- [11] M. Sharp and D. Corp, "ZEPATIER drug label," https://www.merck.com/product/usa/pi_circulars/z/zepatier/zepatier_pi.pdf.
- [12] A. V. Inc, "VIEKIRA PAK drug label," https://www.accessdata.fda.gov/drugsatfda_docs/label/2014/206619lbl.pdf.
- [13] E. K. Fromme, K. M. Eilers, M. Mori, Y. C. Hsieh, and T. M. Beer, "How accurate is clinician reporting of chemotherapy adverse effects? A comparison with patient-reported symptoms from the quality-of-life questionnaire C30," *Journal of Clinical Oncology*, vol. 22, no. 17, pp. 3485–3490, 2004.
- [14] E. Basch, A. Iasonos, T. McDonough et al., "Patient versus clinician symptom reporting using the National Cancer Institute Common Terminology Criteria for Adverse Events: results of a questionnaire-based study," *The Lancet Oncology*, vol. 7, no. 11, pp. 903–909, 2006.
- [15] E. Basch, "The missing voice of patients in drug-safety reporting," *The New England Journal of Medicine*, vol. 362, no. 10, pp. 865–869, 2010.
- [16] C. Xiao, R. Polomano, and D. W. Bruner, "Comparison between patient-reported and clinician-observed symptoms in oncology," *Cancer Nursing*, vol. 36, no. 6, pp. E1–E16, 2013.
- [17] F. Efficace, G. Rosti, N. Aaronson et al., "Patient-versus physician-reporting of symptoms and health status in chronic myeloid leukemia," *Haematologica*, vol. 99, no. 4, pp. 788–793, 2014.
- [18] T. M. Atkinson, A. C. Dueck, D. V. Satele et al., "Clinician vs patient reporting of baseline and postbaseline symptoms for adverse event assessment in cancer clinical trials," *JAMA Oncology*, vol. 6, no. 3, pp. 437–439, 2020.
- [19] E. Basch, B. B. Reeve, S. A. Mitchell et al., "Development of the National Cancer Institute's patient-reported outcomes version of the common terminology criteria for adverse events (PRO-CTCAE)," *Journal of the National Cancer Institute*, vol. 106, no. 9, article dju244, 2014.
- [20] B. B. Reeve, M. McFatrigh, J. W. Mack et al., "Validity and reliability of the pediatric patient-reported outcomes version of the common terminology criteria for adverse events," *JNCI: Journal of the National Cancer Institute*, vol. 112, no. 11, pp. 1143–1152, 2020.
- [21] United States Food & Drug Administration, "Project patient voice," <https://www.fda.gov/about-fda/oncology-center-excellence/project-patient-voice>.
- [22] M. S. Sulkowski, J. S. Moon, K. E. Sherman et al., "A pragmatic, randomized controlled trial of oral antivirals for the treatment of chronic hepatitis C: the PRIORITIZE study," *Hepatology*, vol. 74, no. 6, pp. 2952–2964, 2021.
- [23] P. Mishra, J. Florian, J. Peter et al., "Public-private partnership: targeting real-world data for hepatitis C direct-acting antivirals," *Gastroenterology*, vol. 153, no. 3, pp. 626–631, 2017.
- [24] C. Christodoulou, D. U. Junghaenel, D. A. DeWalt, N. Rothrock, and A. A. Stone, "Cognitive interviewing in the evaluation of fatigue items: results from the patient-reported outcomes measurement information system (PROMIS)," *Quality of Life Research*, vol. 17, no. 10, pp. 1239–1246, 2008.
- [25] D. U. Junghaenel, C. Christodoulou, J. S. Lai, and A. A. Stone, "Demographic correlates of fatigue in the US general population: results from the patient-reported outcomes measurement information system (PROMIS) initiative," *Journal of Psychosomatic Research*, vol. 71, no. 3, pp. 117–123, 2011.
- [26] D. Cella, W. Riley, A. Stone et al., "The Patient-Reported Outcomes Measurement Information System (PROMIS) developed and tested its first wave of adult self-reported health outcome item banks: 2005–2008," *Journal of Clinical Epidemiology*, vol. 63, no. 11, pp. 1179–1194, 2010, [PubMed: 20685078].
- [27] B. B. Reeve, R. D. Hays, J. B. Bjorner et al., "Psychometric evaluation and calibration of health-related quality of life item banks: plans for the patient-reported outcomes measurement information system (PROMIS)," *Medical Care*, vol. 45, no. 5, pp. S22–S31, 2007, [PubMed: 17443115].
- [28] B. M. Spiegel, R. D. Hays, R. Bolus et al., "Development of the NIH patient-reported outcomes measurement information system (PROMIS) gastrointestinal symptom scales,"

- The American Journal of Gastroenterology*, vol. 109, no. 11, pp. 1804–1814, 2014.
- [29] M. Kosinski, M. S. Bayliss, J. B. Bjorner et al., “A six-item short-form survey for measuring headache impact: the HIT-6,” *Quality of Life Research: an International Journal of Quality of Life Aspects of Treatment, Care and Rehabilitation*, vol. 12, no. 8, pp. 963–974, 2003.
 - [30] J. B. Bjorner, M. Kosinski, and J. E. Ware Jr., “Using item response theory to calibrate the Headache Impact Test (HIT™) to the metric of traditional headache scales,” *Quality of Life Research*, vol. 12, no. 8, pp. 981–1002, 2003.
 - [31] D. M. Evon, J. Amador, P. Stewart et al., “Psychometric properties of the PROMIS short form measures in a U.S. cohort of 961 patients with chronic hepatitis C prescribed direct acting antiviral therapy,” *Alimentary Pharmacology & Therapeutics*, vol. 47, no. 7, pp. 1001–1011, 2018.
 - [32] J. R. Landis and G. G. Koch, “The measurement of observer agreement for categorical data,” *Biometrics*, vol. 1, pp. 159–174, 1977.
 - [33] A. A. Neath and J. E. Cavanaugh, “The Bayesian information criterion: background, derivation, and applications,” *Wiley Interdisciplinary Reviews: Computational Statistics*, vol. 4, no. 2, pp. 199–203, 2012.
 - [34] R. Tibshirani, “Regression shrinkage and selection via the lasso,” *Journal of the Royal Statistical Society: Series B (Methodological)*, vol. 58, no. 1, pp. 267–288, 1996.
 - [35] E. McColl, O. Junghard, I. Wiklund, and D. A. Revicki, “Assessing symptoms in gastroesophageal reflux disease: how well do clinicians' assessments agree with those of their patients?,” *American Journal of Gastroenterology*, vol. 100, no. 1, pp. 11–18, 2005.
 - [36] A. M. Barbara, M. Loeb, L. Dolovich, K. Brazil, and M. Russell, “Agreement between self-report and medical records on signs and symptoms of respiratory illness,” *Primary Care Respiratory Journal*, vol. 21, no. 2, pp. 145–152, 2012.
 - [37] A. C. Justice, C. H. Chang, L. Rabeneck, and R. Zackin, “Clinical importance of provider-reported HIV symptoms compared with patient-report,” *Medical Care*, vol. 39, no. 4, pp. 397–408, 2001.
 - [38] B. M. Peck, “Age-related differences in doctor-patient interaction and patient satisfaction,” *Current Gerontology and Geriatrics Research*, vol. 2011, Article ID 137492, 10 pages, 2011.