

Original Paper

# Assessing Physician and Patient Agreement on Whether Patient Outcomes Captured in Clinical Progress Notes Reflect Treatment Success: Cross-Sectional Study

Sarah B Floyd<sup>1</sup>, MPH, PhD; Jordyn C Sutton<sup>1</sup>, BS; Marvin Okon<sup>1</sup>, MD; Mary McCarthy<sup>2</sup>, BA; Liza Fisher<sup>2</sup>, MBA; Benjamin Judkins<sup>3</sup>, MD; Zachary Cole Reynolds<sup>3</sup>, MD; Ann Blair Kennedy<sup>4</sup>, DrPH

<sup>1</sup>Department of Public Health Sciences, Clemson University, Clemson, SC, United States

<sup>2</sup>Patient Engagement Studio, University of South Carolina, Greenville, SC, United States

<sup>3</sup>Department of Orthopaedic Surgery, Prisma Health, Greenville, SC, United States

<sup>4</sup>School of Medicine, University of South Carolina, Greenville, SC, United States

**Corresponding Author:**

Sarah B Floyd, MPH, PhD

Department of Public Health Sciences

Clemson University

516 Edwards Hall

Clemson, SC, 29634

United States

Phone: 1 864 656 6946

Email: [sbf@clemson.edu](mailto:sbf@clemson.edu)

## Abstract

**Background:** It remains unclear if there is agreement between physicians and patients on the definition of treatment success following orthopedic treatment. Clinical progress notes are generated during each health care encounter and include information on current disease symptoms, rehabilitation progress, and treatment outcomes.

**Objective:** This study aims to assess if physicians and patients agree on whether patient outcomes captured in clinical progress notes reflect a successful treatment outcome following orthopedic care.

**Methods:** We performed a cross-sectional analysis of a subset of clinical notes for patients presenting to a Level-1 Trauma Center and Regional Health System for follow-up for an acute proximal humerus fracture (PHF). This study was part of a larger study of 1000 patients with PHF receiving initial treatment between 2019 and 2021. From the full dataset of 1000 physician-labeled notes, a stratified random sample of 25 notes from each outcome label group was identified for this study. A group of 2 patients then reviewed the sample of 100 clinical notes and labeled each note as reflecting treatment success or failure. Cohen  $\kappa$  statistics were used to assess the degree of agreement between physicians and patients on clinical note content.

**Results:** The average age of the patients in the sample was 67 (SD 13) years and 82% of the notes came from female patients. Patients were primarily White (91%) and had Medicare insurance coverage (65%). The note sample came from fracture-related encounters ranging from the second to the tenth encounter after the index PHF visit. There were no significant differences in patient or visit characteristics across concordant and discordant notes labeled by physicians and patients. Among agreement levels ranging from poor to perfect agreement, physician and patient evaluators exhibited only a fair level of agreement in what they deemed as treatment success based on a Cohen  $\kappa$  of 0.32 (95% CI 0.10-0.55;  $P=.01$ ). Furthermore, interpatient and interphysician agreement also demonstrated relatively low levels of agreement.

**Conclusions:** The findings suggest that physicians and patients demonstrated low levels of agreement when assessing whether a patient's clinical note reflected a successful outcome following treatment for a PHF. As low levels of agreement were also observed within physician and patient groups, it is clear the definition of success varied highly across both physicians and patients. Further research is needed to elucidate physician and patient perceptions of treatment success. As outcome measurement and demonstrating the value of orthopedic treatment remain important priorities, it is important to better define and reach a consensus on what treatment success means in orthopedic medicine.

(*J Particip Med* 2025;17:e60263) doi: [10.2196/60263](https://doi.org/10.2196/60263)

**KEYWORDS**

patient outcomes; proximal humerus fracture; patient involvement; orthopaedic medicine; clinical progress notes

## Introduction

In 1910, Ernest Amory Codman, an orthopedic surgeon, advocated for the concept of studying the “end result,” or the idea that every surgeon should follow patients long enough to evaluate whether the treatment they received was successful [1]. Early on, as surgeons began adopting Codman’s end result approach, physician-reported measurement of individual patient outcomes (eg, mortality, surgical complications, and degrees of range of motion) became the standard method to evaluate the success of orthopedic treatment. However, since that time, health care has continued to increase its appreciation of the patient’s perspective on outcome achievement, and patient preferences for outcomes following care [2-6]. As outcome measurement and demonstrating the value of orthopedic treatment are becoming an increasing priority [7,8], it is important to better elucidate what treatment success means in orthopedic medicine [9,10]. To date, it remains unclear if physicians and patients share the same definition of treatment success following orthopedic care.

The electronic health record (EHR) system is the primary tool to document and store records of patient encounters in hospitals and outpatient clinics in the United States [11-13]. Clinical progress notes are generated for each encounter that patients have with their physician or health care provider. These contain rich information on current disease symptoms, rehabilitation progress, and unexpected complications [14]. Unstructured progress notes produce a record of a patient’s history, physical findings, medical reasoning, and patient care and reveal distinct trajectories of patient outcomes after treatment [13,15,16]. In successful cases, the progress note documents the degree of improvement or relief experienced and reported by patients [17]. Conversely, when symptoms have not been resolved, are lingering, or when subsequent complications have arisen, these ongoing patient complaints and persistent treatment use are documented in the notes [18]. Clinical progress notes offer an opportunity to assess a range of outcome states and evaluate if physicians and patients have similar definitions of success following medical treatment for an orthopedic condition. Furthermore, the secondary use of EHR data is rapidly expanding, including the use of natural language processing and large language models to analyze unstructured clinical text [19-25]. One potential application of these methods includes using clinical notes as a data source to evaluate the success of orthopedic treatment. However, to correctly apply this method, a gold-standard definition of treatment success must be identified.

The objective of this paper was to assess agreement between patients and physicians on whether patient outcomes documented in clinical progress notes reflected successful or unsuccessful treatment outcomes for patients receiving follow-up care for a leading shoulder condition, an acute proximal humerus fracture (PHF).

## Methods

### Study Sample

This was a cross-sectional analysis of a subset of progress notes from a larger study. The study included adult patients presenting in person to a Level-1 Trauma Center and Regional Health System for an acute PHF between January 1, 2019, and December 31, 2021. The index visit was defined as the first diagnosis at any health system site for PHF during the study period, with no previous visits for PHF within a year of the index visit. We identified all health system encounters (hospital encounters, office visits, etc) with a diagnosis of PHF or shoulder pain from the index PHF visit to 365 days after the index PHF visit. Of those encounters, we took the progress note from the last in-person office visit for PHF-related care, defined as a visit with a diagnosis of PHF (International Statistical Classification of Diseases and Related Health Problems 10th Revision [ICD10]: S42.2XXX) or shoulder pain (ICD10: M45.2XXX) to occur before 365 days postindex. This resulted in 1 note per person.

Patients were excluded from the study if they were less than 18 years of age, did not have at least 1 office visit with a diagnosis of PHF or shoulder pain that occurred 45 days or more days after the index visit, or if their last office visit was less than 500 characters. A minimum of 45 days after the index was used as this is the minimal time needed for healing of a PHF, before which treatment success cannot be assessed. The larger study included a sample of 1000 patients meeting these inclusion criteria. For this study, a sample of 100 progress notes was used to assess agreement between physicians and patients on their perceptions of treatment outcomes captured in the clinical notes. This study was approved by the Prisma Health Institutional Review Board (1924627-1).

### Outcome Label Development Process

The University of South Carolina Patient Engagement Studio (PES) brings together patients, caregivers, community groups, health system innovators, clinicians, and academic researchers to produce meaningful research that advances health outcomes. The PES membership includes over 100 patients with diverse backgrounds and clinical experiences from across the United States trained to provide feedback and collaborate with research teams [26-28]. PES staff members assembled a panel of 5 patients all of whom had a previous orthopedic experience including a joint injury of the shoulder, wrist, or ankle. These patients experienced a mix of surgical and nonsurgical management for their condition. Specific demographics of the panel are not shared per PES policy as these patients are consultants rather than study participants. PES staff members facilitated the senior author (SBF) to lead 3 sessions to codevelop a range of outcome states following orthopedic treatment. Together, the PES members and senior author defined 4 distinct outcome states that spanned the range of outcomes patients could experience following treatment for PHF.

Figure 1 contains the 4 distinct outcome states, associated definitions, and indicators. The 4 outcome states included “Treatment Success” which is defined as patients resuming desired activities, achieving a sufficient range of motion, and no more than minimal or mild pain; “Improvement of Condition” included cases where there was a record of some level of pain or functional problems, but improvement of the condition was occurring; “Deterioration of Condition” occurred when there

was a record of some level of pain or functional problems that were becoming more prohibitive to the patient’s desired activities and no improvement was occurring; and “Treatment Failure” occurred when the patient was experiencing significant pain or limitations and required subsequent fracture-related care for fracture sequelae, complications, or nonunion. These 4 outcome state labels were available to patients and physician evaluators when labeling each note.

**Figure 1.** Treatment Outcome States, Definitions and Indicators Developed by Patient Engagement Studio and Research Team Members.

<i>Outcome state</i>	<i>Definition</i>	<i>Example indicators of outcome state found in the clinical note</i>
<b>Treatment success</b>	Treatment success occurs when a patient can resume desired activities, has a sufficient range of motion, and is in minimal/mild or no pain. After PHF it is possible for there to be some lingering motion limitations (patient may never return to 100%) or minimal pain, but these issues should not require ongoing treatment or be prohibitive to their desired lifestyle or daily activities.	<ul style="list-style-type: none"> <li>• Radiographic healing noted on x-ray</li> <li>• Making good progress/improvements with current treatment or stopping treatment</li> <li>• Patient has returned to work or play.</li> <li>• No major complaints documented.</li> <li>• Only follow-up as needed</li> </ul>
<b>Improvement of condition</b>	Improvement occurs when there is a record of some levels of pain or functional problems that are somewhat prohibitive to the patient’s desired activities, but improvement is occurring. In these situations, physicians may continue to monitor patients, but do not alter care or treatment courses.	<ul style="list-style-type: none"> <li>• Radiographic healing or signs of healing occurring.</li> <li>• Moderate loss of function or pain that interferes with desired activities, but no change in treatment.</li> <li>• Ongoing treatment and monitoring progress</li> <li>• Return in 2-6 weeks for repeat x-rays and recheck</li> </ul>
<b>Deterioration of condition</b>	Deterioration occurs when there is a record of some levels of pain or functional problems that are becoming more prohibitive to the patient’s desired activities. No real improvements are occurring, and physicians may escalate or alter care or treatment courses.	<ul style="list-style-type: none"> <li>• Negative radiographic changes observed.</li> <li>• Moderate loss of function or pain that interferes with desired activities requiring a change in treatment.</li> <li>• Initiating or continuing treatment and monitoring progress.</li> <li>• Return in 2-6 weeks for repeat x-rays and recheck.</li> </ul>
<b>Treatment failure</b>	Treatment failure occurs when the patient is experiencing significant pain or limitations and requires subsequent fracture-related care. Failing occurs when patients are unable to resume desired activities and may include fracture sequelae, complications, or nonunion.	<ul style="list-style-type: none"> <li>• Ongoing, persistent treatment (injections, surgeries) for symptoms related to PHF.</li> <li>• Unrelenting pain</li> <li>• Surgical complications</li> <li>• Loss of significant motion</li> <li>• Extreme pain</li> <li>• Fracture-related sequelae (eg, avascular necrosis)</li> </ul>

**Note Labeling Process**

*Physician Evaluators*

A total of 4 orthopedic residents were recruited to participate in the note-labeling process as part of the larger study. Each orthopedic resident received a 1-hour training on the study objective and outcome state labels. Residents were instructed to assess the current outcome state reflected in the note. The

physician evaluators included 3 male and 1 female orthopedic residents, each of which had a minimum of 2 years of residency experience. When discordance occurred between residents’ labels, an attending orthopedic surgeon and the Chair of the Department of Orthopaedic Surgery served as the final note evaluator. REDCap (Research Electronic Data Capture; Vanderbilt University) [29,30] was used to organize and store physician labels for each note. From the full dataset of 1000 labeled notes, a stratified random sample of 25 notes from each

outcome label group was identified, and the note sample (N=100) for patient labeling was created.

### **Patient Evaluators**

We recruited 2 patients from the PES to participate in this study. Both patients were female and had personal orthopedic experience including upper and lower extremity conditions, but their personal clinical data were not included in our study sample. The patient evaluators brought both experiential expertise from their personal musculoskeletal conditions and specialized research training, enabling them to contribute effectively to this study. This aligns with current best practices in patient engagement, which emphasize the value of relevant patient perspectives and training over the necessity for identical clinical conditions [31-34]. Similar to the physician evaluators, patient evaluators also received a 1-hour training on the study objective and outcome state labels. The training included a group review of example charts and common language used in medical charts. In addition, we trained patients in the subjective, objective, assessment, and plan sections [14] format typically used in medical documentation to increase their familiarity with navigating a medical chart. All clinical progress notes were redacted to conceal patient identifiers before patient review.

Both patient evaluators reviewed all 100 notes and provided labels. In addition to the 4 outcome state labels, a label of “Insufficient” was available for patient evaluators for notes deemed to have insufficient information to assign an outcome label. When discordance occurred between patient evaluators, the Program Manager of the PES (KP) served as the final note evaluator. After review by the Program Manager, all notes had a final label, and all labels of “insufficient” were resolved.

### **Patient and Visit Characteristics**

Patient characteristics associated with the 100 clinical notes included in the analysis were extracted from the health system EHR, Epic, and included patient age, sex, race, and insurance provider. Patient characteristics were identified from the index PHF visit. In addition, visit characteristics, including days between the index visit and visit date for the clinical note, the number of PHF-related encounters, surgical treatment use, and note length, were also included in the analysis. Patients receiving surgery were defined as those patients undergoing reverse shoulder arthroplasty, hemiarthroplasty, or open reduction internal fixation between the index and 365 days.

### **Statistical Analysis**

The 4 outcome labels were aggregated into a binary classifier representing treatment success or failure. Success was represented by notes labeled “Treatment Success.” The 3 remaining labels, including “Improvement of condition,” “Deterioration of condition,” and “Treatment Failure” were grouped into the Treatment failure group. Treatment failure was comprised of all labels with documentation of lingering, symptomatic problems requiring ongoing care.

Agreement between physicians and patients was calculated across binary groups of treatment success or failure. Discordant labels were defined as notes with differing outcome states provided by the respective labelers. Cohen  $\kappa$  statistics were used to assess the degree of agreement between patient evaluators, as well as the degree of agreement between physician and patient labels. In addition, physician agreement was reported for the larger sample of 1000 notes and was assessed using Fleiss  $\kappa$  [35]. We used the benchmarks for agreement for categorical data as described by Landis and Koch [36], where 0.00-0.20, 0.21-0.40, 0.41-0.60, 0.61-0.80, and 0.81-1.00 indicate poor, fair, moderate, substantial, and almost perfect agreement, respectively. A Bangdiwala agreement chart is presented to display the agreement between physician and patient labels [37].

Descriptive analyses were used to assess the characteristics of the progress note sample. Mean and SD were reported for parametric variables. Median and IQR (25% and 75%) were reported for nonparametric variables. Two-sample  $t$  test, Wilcoxon-Mann-Whitney, and chi-square tests were used to assess differences in concordant and discordant notes. Analyses were performed with SAS (version 15.2; SAS Institute), R studio (R Core Team), and Microsoft Excel.

## **Results**

### **Progress Note Characteristics**

The sample of 100 progress notes for this study came from patients treated across 24 departments and 54 distinct physicians within one regional health system. The 24 departments from which the notes were identified included 21 orthopedic practices or departments, 2 family medicine, and 1 pain management clinic. Notes were authored by both physicians and advanced practice providers. Of the 41 physicians, 35 (85%) specialized in orthopedics, whereas the remaining 6 (15%) were specialists in family medicine. In addition to the 41 physicians, 13 advanced practice providers completed notes and 10 (77%) of these providers specialized in orthopedics, while the remainder had other specialty training including general surgery and pain medicine.

The average age of the patient was 67 (SD 13) years and 82% of the notes came from female patients. Patients were primarily White (91%) and had Medicare insurance coverage (65%). The note sample came from fracture-related encounters ranging from the second to the 10th encounter after the index PHF visit, with a median time of 115 (IQR 73-215) days after the index. The progress notes text lengths ranged from 981 to 15,297 characters with a median length of 5098 (IQR 2846-7810) characters. There was no significant difference in progress note characteristics across concordant and discordant notes (Table 1).

**Table 1.** Patient and visit characteristics of the clinical progress note sample presented by patient and physician agreement (N=100). Mean and SD were reported for parametric variables. Median and IQR (25% and 75%) are reported for nonparametric variables. A 2-sample *t* test was used for parametric variables and the Wilcoxon-Mann-Whitney test was used for nonparametric comparisons.

Patient characteristics	Total sample (N=100)	Concordant notes (n=78)	Discordant notes (n=22)	<i>P</i> value
Patient age (years), mean (SD)	67 (13)	67 (13)	68 (13)	.73
<b>Patient sex, n (%)</b>				.22
Male	18 (18)	16 (20)	2 (9)	— <sup>a</sup>
Female	82 (82)	62 (79)	20 (90)	—
<b>Patient race, n (%)</b>				.72
White	91 (91)	71 (91)	20 (91)	—
Black	5 (5)	3 (4)	2 (9)	—
American Indian or Alaskan	1 (1)	1 (1)	0 (0)	—
Hispanic	1 (1)	1 (1)	0 (0)	—
Unknown	2 (2)	2 (3)	0 (0)	—
<b>Insurance provider, n (%)</b>				
Medicare	65 (65)	51 (65)	14 (64)	.44
Medicaid	7 (7)	7 (9)	0 (0)	—
Private	21 (21)	15 (19)	6 (27)	—
Other	7 (7)	5 (6)	2 (9)	—
<b>Visit characteristics</b>				
Days from index, median (IQR)	115 (73-215)	113 (74-219)	115 (65-170)	.65
PHF <sup>b</sup> -related encounter, median (IQR)	4 (3-6)	4 (3-6)	4 (3-6)	.44
Patient treated surgically, n (%)	25 (25)	21 (27)	4 (18)	.40
Note character length, median (IQR)	5098 (2846-7810)	5202 (2901-8155)	4320 (2672-6428)	.19

<sup>a</sup>Not applicable.

<sup>b</sup>PHF: proximal humerus fracture.

### Agreement Between Patients

Both patient evaluators were assigned the full sample of 100 notes to review and label. Of the 100 notes, 34 notes were discordant between patient evaluators. A total of 23 of the discordant labels were between success and failure labels between patient evaluators. In addition, there were a total of 11

cases (across patient evaluators 1 and 2) that received a label of “insufficient.” There was a statistically significant level of agreement between the 2 patient evaluators (Cohen  $\kappa=0.41$ , 95% CI 0.23-0.59;  $P<.001$ ), and the strength of agreement was classified as moderate, according to Landis and Koch. [Tables 2 and 3](#) show the agreement in note labels between patient evaluators and physicians and patient evaluators.

**Table 2.** Agreement in note labels between patients (N=100).

Patient rater 1	Patient rater 2			Total	Agreement
	Success	Failure	Indeterminate <sup>a</sup>		
Success	15	3	1	19	Moderate ( $\kappa=0.41$ ) <sup>b</sup>
Failure	20	51	8	79	
Indeterminate <sup>a</sup>	0	2	0	2	
Total	35	56	9	100	

<sup>a</sup>A label of indeterminate was available for use by patient evaluators for notes deemed to have insufficient information for a label. Notes labeled as insufficient were reviewed by the PES Manager for final label assignment. After final review, all notes had a final label, and all labels of insufficient were resolved before future analysis.

<sup>b</sup>Cohen  $\kappa$  used to assess agreement. 0.00-0.20, 0.21-0.40, 0.41-0.60, 0.61-0.80, and 0.81-1.00 indicate poor, fair, moderate, substantial, and almost perfect agreement.

**Table 3.** Agreement in note labels between physicians and patients (N=100).

Physician labels	Patient labels			Agreement
	Success	Failure	Total	
Success	11	14	25	Fair ( $\kappa=0.32$ ) <sup>a</sup>
Failure	8	67	75	
Total	19	81	100	

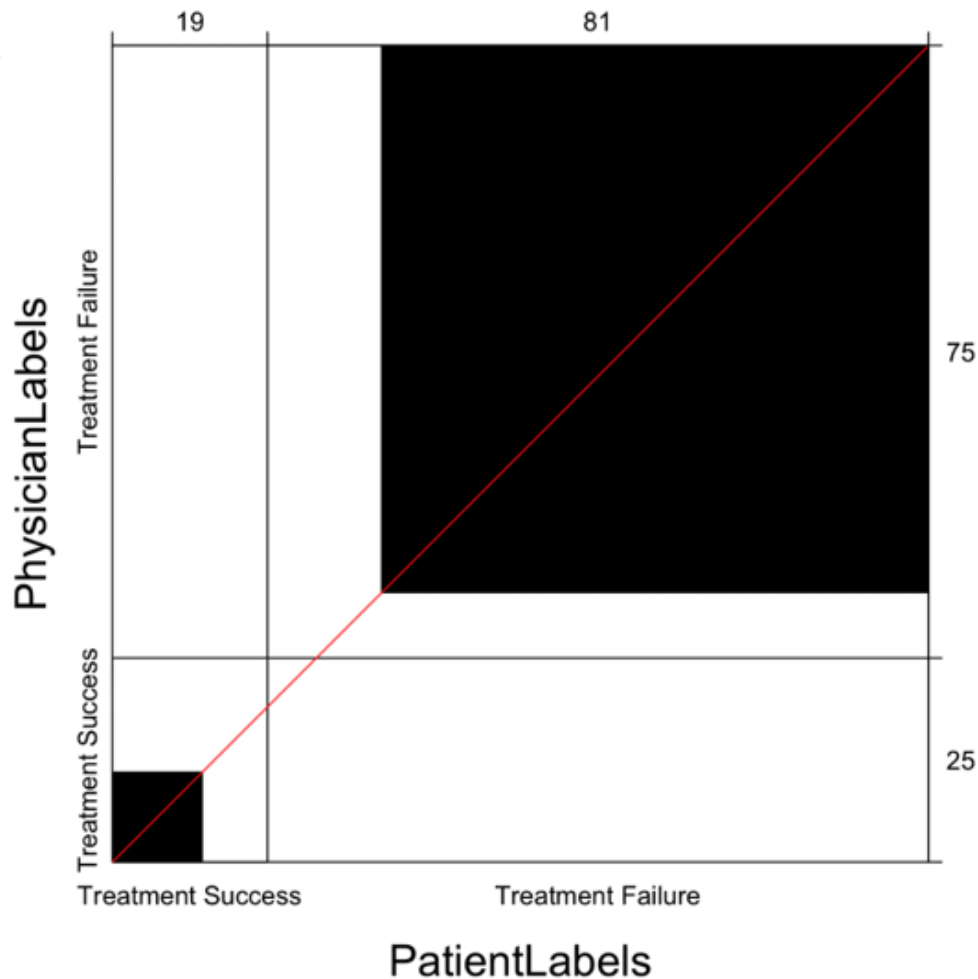
<sup>a</sup>Cohen  $\kappa$  used to assess agreement. 0.00-0.20, 0.21-0.40, 0.41-0.60, 0.61-0.80, and 0.81-1.00 indicate poor, fair, moderate, substantial, and almost perfect agreement.

**Agreement Between Physicians and Patients**

A total of 22 notes were discordant between physicians and patient evaluators. Of the 25 notes labeled as treatment success by orthopedic surgeons, 11 notes were also labeled as treatment success by patients. The remaining 14 treatment success notes were labeled as treatment failure by patient evaluators. Of the 75 notes deemed as treatment failure, 67 were also labeled as

treatment failure by patient evaluators. There was a statistically significant level of agreement between orthopedic physicians and patient evaluators (Cohen  $\kappa=0.32$ , 95% CI 0.10-0.55;  $P=.01$ ). The strength of agreement between patients and physicians was classified as fair, according to Landis and Koch. Figure 2 includes a Bangdiwala chart used to display agreement between patients’ and physicians’ assessment of treatment success or treatment failure from analyzed clinical notes.

**Figure 2.** Bangdiwala agreement chart for physician and patient note labels (N=100). Bangdiwala chart used to assess agreement between patients and physician’s indications of treatment success or treatment failure from analyzed clinical notes. Black boxes indicate overlap of agreement.



Although not the focus of this paper, physician agreement was assessed using the larger sample of 1000 notes. Agreement between physicians was assessed using Fleiss  $\kappa$  and agreement between orthopedic physicians was moderate (Fleiss  $\kappa=0.49$ , 95% CI 0.30-0.68;  $P=.04$ ).

**Discussion**

**Principal Findings**

The objective of this paper was to assess if physicians and patients agree in their assessment of whether patient outcomes

in clinical progress notes reflected a successful treatment outcome following orthopedic care. This is an important question to answer for the field of orthopedic medicine which has experienced a paradigm shift in the way in which outcomes are assessed [3,38,39]. Outcome assessment in orthopedics dates back over 100 years. Early on, physician-reported measurement of individual patient outcomes was the standard method by which to evaluate the outcomes of orthopedic care. However, today outcome measurement directly from a patient's perspective is viewed as the gold standard in orthopedic medicine [39,40]. We were interested in exploring if patients and physicians have similar definitions of what successful outcomes mean following orthopedic treatment.

In our analysis, we had patients and physicians review a subset of 100 clinical progress notes and label the note as a successful or unsuccessful outcome. We found that physicians and patients only exhibited a fair level of agreement in what they deemed as treatment success documented in progress notes. In addition, we found that physicians and patients had higher levels of agreement in what represented treatment failure compared with treatment success. Furthermore, interpatient and interphysician agreements also demonstrated relatively low levels of agreement, signaling that even within patients and physician groups, the definition of success is not clearly defined or agreed upon.

### Comparison to Previous Work

A potential explanation for the low level of agreement between patients and physicians may simply be that patients and physicians have different expectations following care. Our findings might signal that physicians have different expectations of patient's capabilities following a serious upper extremity injury, such as PHF [41,42]. For other orthopedic treatments, it has been reported that patient expectations may be greater than a physician's expectations [43]. For example, in total hip and knee arthroplasty, most patients had higher expectations for recovery than their surgeon [43]. This might explain why over half of the notes labeled as treatment success by orthopedic surgeons were labeled as treatment failure by patients. Patients appeared to have a more stringent definition of success compared with physicians. Although not the goal of our study, this finding does emphasize the importance of shared decision-making within orthopedic encounters, to ensure patients have realistic expectations of outcomes following care [44].

An alternative explanation for our finding could be that physicians and patients define success differently. In a study assessing patient-physician agreement on the management of musculoskeletal injuries and pain associated with those injuries, authors found that patients and physicians prioritize different goals when assessing a patient's treatment outcome [4,45]. For example, physicians may have a more clinically based definition of treatment success driven by objective measures such as radiographic measures of healing and degrees of range of motion, whereas patients may be more focused on the ways in which outcomes like pain and joint function relate to daily capabilities and quality of life [5].

We found that physicians and patients had higher levels of agreement in what represented treatment failure compared with

treatment success. Other studies measuring patient and physician agreement following orthopedic surgery concluded that patients and physicians agreed more when the patient had good health outcomes [4,46,47]. These conclusions are not consistent with our study findings. We found that physicians and patients were in agreement for a larger share of the treatment failure notes, compared with the treatment success notes. It is our belief that treatment failure is more clear-cut (eg, surgical complications, persistent pain, and fracture nonunion), whereas treatment success is more variable and patient-specific. Consequently, it may be easier to recognize when outcomes are unfavorable, but pinpointing a positive outcome proves challenging due to the variability and outcome preferences across individual patients [48,49]. Furthermore, we believe the concept of a patient-specific definition of success is supported by the moderate level of agreement we observed between patients. This signals that even among patients, there is a differential evaluation of an acceptable outcome. There is not 1 singular definition of treatment success, instead, treatment success depends on an individual patient's lifestyle and desired goals. Finally, even among physicians, we still observed relatively low levels of agreement, signaling that the definition of success remains unclear across physicians.

### Limitations

Our work has several limitations that should be acknowledged. First, we used a relatively small sample of progress notes from 1 clinical condition that lacks patient diversity. Furthermore, our results are highly reflective of the small sample of physicians and patient evaluators who completed the labeling. Next, we were unable to assess the characteristics of treating physicians who authored the progress notes. It is possible physician characteristics like subspecialty training, years of experience, and so on, may explain some of the discordance in note labels. In addition, we worked with resident physicians who may be less experienced in assessing patient outcomes following care. This could affect physician agreement, as well as physician-patient agreement results. Also, the way in which we aggregated patient labels may influence the level of agreement we observed. For example, more categories could potentially lead to lower concordance among evaluators. Finally, it is possible that as nonmedically trained individuals patient evaluators' labeling may have been influenced by their lack of medical training.

### Future Directions

Although outside the scope of this work, there remain questions surrounding the accuracy of clinical notes. There are mixed reports of the accuracy, completeness, and quality of progress note content [50-53]. Multiple studies have found that health care professionals produce accurate documentation for concrete and overt symptoms, such as range of motion and impaired physical functioning [54]. However, it must be acknowledged that we did not directly assess the accuracy of physician reporting of patient outcomes captured in the clinical notes. Secondary use of EHR data is rapidly expanding, including the use of natural language processing and large language models to analyze unstructured clinical text [19-25]. One potential use could be to use clinical notes to evaluate the success of

orthopedic treatment. However, to appropriately assess and classify outcomes as either successful or unsuccessful, the accuracy of clinical notes must be assessed.

In addition, as we work to continue to understand the concept of treatment success in orthopedic medicine, it may be helpful to conduct follow-up interviews with physicians and patients as they conclude the labeling process. This could reveal a deeper understanding of each perspective on what treatment success means. Furthermore, we anticipate that future work will incorporate multiple clinical notes across the episode of care to capture a more complete outcome assessment, as interim visits may reveal incremental improvements before the final visit.

## Conclusion

The objective of this study was to assess if physicians and patients agree on whether patient experiences captured in clinical progress notes reflect a successful patient outcome following orthopedic treatment. In performing a cross-sectional analysis of clinical progress notes from an acute follow-up of patients treated for a PHF, we found fair agreement between patients' and physicians' assessments of patient outcomes reflecting treatment success. These results indicate that patients and physicians do not fully agree on what constitutes treatment success. Our findings emphasize the need to analyze both patient and physician perspectives when determining treatment success. Further research is needed to examine how different perceptions of treatment success may influence outcome development and use in orthopedic medicine.

## Acknowledgments

We want to acknowledge the contributions of the orthopedic residents who completed the note labeling. We thank Drs. Kyle Jeray, Jahan Threeths, and Claire Krohn. We also thank Katie Parris, the Patient Engagement Studio Program Manager, for her partnership in completing the patient labels. This study was funded by the Agency for Healthcare Research and Quality (R03HS029060). No generative AI was used in any portion of manuscript generation.

## Data Availability

The datasets generated during and/or analyzed during this study are available from the corresponding author on reasonable request.

## Authors' Contributions

SF and ABK contributed to study conceptualization, data analysis, result interpretation, and manuscript drafting and editing. JS, MO, MM, LF, BJ, and ZR handled data analysis, result interpretation, and manuscript drafting and editing.

## Conflicts of Interest

None declared.

## References

1. Nakayama DK. Ernest Amory Codman and the end result idea in surgical quality. *Am Surg*. 2023;89(11):4237-4240. [doi: [10.1177/00031348221148344](https://doi.org/10.1177/00031348221148344)] [Medline: [36565264](https://pubmed.ncbi.nlm.nih.gov/36565264/)]
2. Shapiro LM, Eppler SL, Roe AK, Morris A, Kamal RN. The patient perspective on patient-reported outcome measures following elective hand surgery: a convergent mixed-methods analysis. *J Hand Surg Am*. Feb 2021;46(2):153.e1-153.e11. [FREE Full text] [doi: [10.1016/j.jhsa.2020.09.008](https://doi.org/10.1016/j.jhsa.2020.09.008)] [Medline: [33183858](https://pubmed.ncbi.nlm.nih.gov/33183858/)]
3. Dean S, Al Sayah F, Johnson JA. Measuring value in healthcare from a patients' perspective. *J Patient Rep Outcomes*. 2021;5(Suppl 2):88. [FREE Full text] [doi: [10.1186/s41687-021-00364-4](https://doi.org/10.1186/s41687-021-00364-4)] [Medline: [34637001](https://pubmed.ncbi.nlm.nih.gov/34637001/)]
4. Noble PC, Fuller-Lafreniere S, Meftah M, Dwyer MK. Challenges in outcome measurement: discrepancies between patient and provider definitions of success. *Clin Orthop Relat Res*. 2013;471(11):3437-3445. [FREE Full text] [doi: [10.1007/s11999-013-3198-x](https://doi.org/10.1007/s11999-013-3198-x)] [Medline: [23955192](https://pubmed.ncbi.nlm.nih.gov/23955192/)]
5. Berlin C, Tieleman S, Quante M, Halm H. Correlation of radiographic parameters and patient satisfaction in adolescent idiopathic scoliosis treated with posterior screw-dual-rod instrumentation. *Eur Spine J*. 2023;32(9):3140-3148. [doi: [10.1007/s00586-023-07849-4](https://doi.org/10.1007/s00586-023-07849-4)] [Medline: [37470846](https://pubmed.ncbi.nlm.nih.gov/37470846/)]
6. Karsdal MA, Tambiah J, Felson D, Ladel C, Nikolov NP, Hodgins D, et al. Reflections from the OARSI 2022 clinical trials symposium: the pain of OA-deconstruction of pain and patient-reported outcome measures for the benefit of patients and clinical trial design. *Osteoarthritis Cartilage*. 2023;31(10):1293-1302. [FREE Full text] [doi: [10.1016/j.joca.2023.06.006](https://doi.org/10.1016/j.joca.2023.06.006)] [Medline: [37380011](https://pubmed.ncbi.nlm.nih.gov/37380011/)]
7. Pasqualini I, Piuzzi NS. New CMS policy on the mandatory collection of patient-reported outcome measures for total hip and knee arthroplasty by 2027: what orthopaedic surgeons should know. *J Bone Joint Surg Am*. 2024;106(13):1233-1241. [doi: [10.2106/JBJS.23.01013](https://doi.org/10.2106/JBJS.23.01013)] [Medline: [38335264](https://pubmed.ncbi.nlm.nih.gov/38335264/)]
8. Attarian DE, Seyler TM, Bolognesi MP. Overview: value-based care for orthopedic surgery. *Techniques in Orthopaedics*. 2019;35(1). [doi: [10.1097/bto.0000000000000415](https://doi.org/10.1097/bto.0000000000000415)]
9. Suk M, Norvell DC, Hanson B, Dettori JR, Helfet D. Evidence-based orthopaedic surgery: what is evidence without the outcomes? *J Am Acad Orthop Surg*. 2008;16(3):123-129. [doi: [10.5435/00124635-200803000-00003](https://doi.org/10.5435/00124635-200803000-00003)] [Medline: [18316710](https://pubmed.ncbi.nlm.nih.gov/18316710/)]



10. Debbi EM, Khilnani T, Gkias I, Chiu Y, Miller AO, Henry MW, et al. Changing the definition of treatment success alters treatment outcomes in periprosthetic joint infection: a systematic review and meta-analysis. *J Bone Jt Infect.* 2024;9(2):127-136. [FREE Full text] [doi: [10.5194/jbji-9-127-2024](https://doi.org/10.5194/jbji-9-127-2024)] [Medline: [38895103](https://pubmed.ncbi.nlm.nih.gov/38895103/)]
11. Adler-Milstein J, Holmgren AJ, Kralovec P, Worzala C, Searcy T, Patel V. Electronic health record adoption in US hospitals: the emergence of a digital 'advanced use' divide. *J Am Med Inform Assoc.* 2017;24(6):1142-1148. [FREE Full text] [doi: [10.1093/jamia/ocx080](https://doi.org/10.1093/jamia/ocx080)] [Medline: [29016973](https://pubmed.ncbi.nlm.nih.gov/29016973/)]
12. Office-based physician electronic health record adoption, health IT quick-stat #50. Office of the National Coordinator for Health Information Technology. 2019. URL: <https://www.healthit.gov/data/quickstats/office-based-physician-electronic-health-record-adoption> [accessed 2024-12-14]
13. Walsh SH. The clinician's perspective on electronic health records and how they can affect patient care. *BMJ.* 2004;328(7449):1184-1187. [FREE Full text] [doi: [10.1136/bmj.328.7449.1184](https://doi.org/10.1136/bmj.328.7449.1184)] [Medline: [15142929](https://pubmed.ncbi.nlm.nih.gov/15142929/)]
14. Podder VL, Lew V, Ghassemzadeh S. SOAP Notes. Treasure Island (FL). StatPearls Publishing; 2021.
15. Boag W, Doss D, Naumann T, Szolovits P. What's in a note? Unpacking predictive value in clinical note representations. *AMIA Jt Summits Transl Sci Proc.* 2018;2017:26-34. [FREE Full text] [Medline: [29888035](https://pubmed.ncbi.nlm.nih.gov/29888035/)]
16. Varpio L, Rashotte J, Day K, King J, Kuziemy C, Parush A. The EHR and building the patient's story: a qualitative investigation of how EHR use obstructs a vital clinical activity. *Int J Med Inform.* 2015;84(12):1019-1028. [doi: [10.1016/j.ijmedinf.2015.09.004](https://doi.org/10.1016/j.ijmedinf.2015.09.004)] [Medline: [26432683](https://pubmed.ncbi.nlm.nih.gov/26432683/)]
17. Resnik P, Niv M, Nossal M, Kapit A, Toren R. Communication of clinically relevant information in electronic health records: a comparison between structured data and unrestricted physician language. *Perspectives in Health Information Managemen.* 2008;2008:1-10.
18. Murff HJ, FitzHenry F, Matheny ME, Gentry N, Kotter KL, Crimin K, et al. Automated identification of postoperative complications within an electronic medical record using natural language processing. *JAMA.* 2011;306(8):848-855. [doi: [10.1001/jama.2011.1204](https://doi.org/10.1001/jama.2011.1204)] [Medline: [21862746](https://pubmed.ncbi.nlm.nih.gov/21862746/)]
19. Allada AK, Wang Y, Jindal V, Babee M, Tizhoosh HR, Crowley M. Analysis of language embeddings for classification of unstructured pathology reports. In: Annual International Conference. IEEE; 2021. Presented at: 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC); 2021 November 1-5:2378-2381; Mexico. [doi: [10.1109/embc46164.2021.9630347](https://doi.org/10.1109/embc46164.2021.9630347)]
20. Bovonratwet P, Shen TS, Islam W, Ast MP, Haas SB, Su EP. Natural language processing of patient-experience comments after primary total knee arthroplasty. *J Arthroplasty.* 2021;36(3):927-934. [doi: [10.1016/j.arth.2020.09.055](https://doi.org/10.1016/j.arth.2020.09.055)] [Medline: [33127238](https://pubmed.ncbi.nlm.nih.gov/33127238/)]
21. Lee J, Yoon W, Kim S, Kim D, Kim S, So CH, et al. BioBERT: a pre-trained biomedical language representation model for biomedical text mining. *Bioinformatics.* 2020;36(4):1234-1240. [FREE Full text] [doi: [10.1093/bioinformatics/btz682](https://doi.org/10.1093/bioinformatics/btz682)] [Medline: [31501885](https://pubmed.ncbi.nlm.nih.gov/31501885/)]
22. Lee RY, Kross EK, Torrence J, Li KS, Sibley J, Cohen T, et al. Assessment of natural language processing of electronic health records to measure goals-of-care discussions as a clinical trial outcome. *JAMA Netw Open.* 2023;6(3):e231204. [FREE Full text] [doi: [10.1001/jamanetworkopen.2023.1204](https://doi.org/10.1001/jamanetworkopen.2023.1204)] [Medline: [36862411](https://pubmed.ncbi.nlm.nih.gov/36862411/)]
23. Velupillai S, Suominen H, Liakata M, Roberts A, Shah AD, Morley K, et al. Using clinical natural language processing for health outcomes research: overview and actionable suggestions for future advances. *J Biomed Inform.* 2018;88:11-19. [FREE Full text] [doi: [10.1016/j.jbi.2018.10.005](https://doi.org/10.1016/j.jbi.2018.10.005)] [Medline: [30368002](https://pubmed.ncbi.nlm.nih.gov/30368002/)]
24. Wyatt JM, Booth GJ, Goldman AH. Natural language processing and its use in orthopaedic research. *Curr Rev Musculoskelet Med.* 2021;14(6):392-396. [FREE Full text] [doi: [10.1007/s12178-021-09734-3](https://doi.org/10.1007/s12178-021-09734-3)] [Medline: [34755276](https://pubmed.ncbi.nlm.nih.gov/34755276/)]
25. Obeid JS, Khalifa A, Xavier B, Bou-Daher H, Rockey DC. An AI approach for identifying patients with cirrhosis. *J Clin Gastroenterol.* 2023;57(1):82-88. [FREE Full text] [doi: [10.1097/MCG.0000000000001586](https://doi.org/10.1097/MCG.0000000000001586)] [Medline: [34238846](https://pubmed.ncbi.nlm.nih.gov/34238846/)]
26. Kennedy AB, Mitcham A, Parris K, Albertson F, Sanchez Ferrer L, O'Boyle C, et al. Wonderings to research questions: engaging patients in long COVID research prioritization within a learning health system. *Learn Health Syst.* 2024;8(Suppl 1):e10410. [FREE Full text] [doi: [10.1002/lrh2.10410](https://doi.org/10.1002/lrh2.10410)] [Medline: [38883877](https://pubmed.ncbi.nlm.nih.gov/38883877/)]
27. Fleming PR, Swygert MM, Hasenkamp C, Sterling J, Cartee G, Russ-Sellers R, et al. Patient engagement in fertility research: bench research, ethics, and social justice. *Res Involv Engagem.* 2021;7(1):29. [FREE Full text] [doi: [10.1186/s40900-021-00278-x](https://doi.org/10.1186/s40900-021-00278-x)] [Medline: [33980313](https://pubmed.ncbi.nlm.nih.gov/33980313/)]
28. Natafagi N, Parris K, Walker E, Gartner T, Coffin J, Mitcham A, et al. Through their eyes: using photovoice to capture the capacity-building journey of long covid patient experts. *Health Expect.* 2024;27(6):e70094. [FREE Full text] [doi: [10.1111/hex.70094](https://doi.org/10.1111/hex.70094)] [Medline: [39529266](https://pubmed.ncbi.nlm.nih.gov/39529266/)]
29. Harris PA, Taylor R, Minor BL, Elliott V, Fernandez M, O'Neal L, et al. REDCap Consortium. The REDCap consortium: building an international community of software platform partners. *J Biomed Inform.* 2019;95:103208. [FREE Full text] [doi: [10.1016/j.jbi.2019.103208](https://doi.org/10.1016/j.jbi.2019.103208)] [Medline: [31078660](https://pubmed.ncbi.nlm.nih.gov/31078660/)]
30. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform.* 2009;42(2):377-381. [FREE Full text] [doi: [10.1016/j.jbi.2008.08.010](https://doi.org/10.1016/j.jbi.2008.08.010)] [Medline: [18929686](https://pubmed.ncbi.nlm.nih.gov/18929686/)]

31. Arumugam A, Phillips LR, Moore A, Kumaran SD, Sampath KK, Migliorini F, et al. Patient and public involvement in research: a review of practical resources for young investigators. *BMC Rheumatol*. 2023;7(1):2. [FREE Full text] [doi: [10.1186/s41927-023-00327-w](https://doi.org/10.1186/s41927-023-00327-w)] [Medline: [36895053](https://pubmed.ncbi.nlm.nih.gov/36895053/)]
32. de Wit M, Aouad K, Elhai M, Benavent D, Bertheussen H, Blackburn S, et al. EULAR recommendations for the involvement of patient research partners in rheumatology research: 2023 update. *Ann Rheum Dis*. 2024;83(11):1443-1453. [doi: [10.1136/ard-2024-225566](https://doi.org/10.1136/ard-2024-225566)] [Medline: [38876509](https://pubmed.ncbi.nlm.nih.gov/38876509/)]
33. Skovlund PC, Nielsen BK, Thaysen HV, Schmidt H, Finset A, Hansen KA, et al. The impact of patient involvement in research: a case study of the planning, conduct and dissemination of a clinical, controlled trial. *Res Involv Engagem*. 2020;6:43. [FREE Full text] [doi: [10.1186/s40900-020-00214-5](https://doi.org/10.1186/s40900-020-00214-5)] [Medline: [32699648](https://pubmed.ncbi.nlm.nih.gov/32699648/)]
34. Strassle CL, Pearson SD. A proposed framework for patient engagement throughout the broader research enterprise. *J Comp Eff Res*. 2020;9(6):387-393. [FREE Full text] [doi: [10.2217/cer-2019-0175](https://doi.org/10.2217/cer-2019-0175)] [Medline: [32253919](https://pubmed.ncbi.nlm.nih.gov/32253919/)]
35. Fleiss JL. Measuring nominal scale agreement among many raters. *Psychological Bulletin*. 1971;76(5):378-382. [doi: [10.1037/h0031619](https://doi.org/10.1037/h0031619)]
36. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977;33(1):159-174. [Medline: [843571](https://pubmed.ncbi.nlm.nih.gov/843571/)]
37. Bangdiwala SI, Shankar V. The agreement chart. *BMC Med Res Methodol*. 2013;13:97. [FREE Full text] [doi: [10.1186/1471-2288-13-97](https://doi.org/10.1186/1471-2288-13-97)] [Medline: [23890315](https://pubmed.ncbi.nlm.nih.gov/23890315/)]
38. Deshpande PR, Rajan S, Sudeepthi BL, Abdul Nazir CP. Patient-reported outcomes: a new era in clinical research. *Perspect Clin Res*. 2011;2(4):137-144. [FREE Full text] [doi: [10.4103/2229-3485.86879](https://doi.org/10.4103/2229-3485.86879)] [Medline: [22145124](https://pubmed.ncbi.nlm.nih.gov/22145124/)]
39. Gagnier JJ. Patient reported outcomes in orthopaedics. *J Orthop Res*. 2017;35(10):2098-2108. [FREE Full text] [doi: [10.1002/jor.23604](https://doi.org/10.1002/jor.23604)] [Medline: [28513993](https://pubmed.ncbi.nlm.nih.gov/28513993/)]
40. Jayakumar P, Heng M, Okelana B, Vrahas M, Rodriguez-Villalon A, Joeris A. Patient-reported outcome measurement in orthopaedic trauma. *J Am Acad Orthop Surg*. 2023;31(20):e906-e919. [doi: [10.5435/JAAOS-D-23-00375](https://doi.org/10.5435/JAAOS-D-23-00375)] [Medline: [37796280](https://pubmed.ncbi.nlm.nih.gov/37796280/)]
41. Dy CJ, Gonzalez Della Valle A, York S, Rodriguez JA, Sculco TP, Ghomrawi HM. Variations in surgeons' recovery expectations for patients undergoing total joint arthroplasty: a survey of the AAHKS membership. *J Arthroplasty*. 2013;28(3):401-405. [doi: [10.1016/j.arth.2012.06.035](https://doi.org/10.1016/j.arth.2012.06.035)] [Medline: [23142456](https://pubmed.ncbi.nlm.nih.gov/23142456/)]
42. Karpinski K, Plachel F, Gerhardt C, Saier T, Tauber M, Auffarth A, et al. Differences in patients' and surgeons' expectations before shoulder stabilization surgery. *J Clin Med*. 2021;10(20):4661. [FREE Full text] [doi: [10.3390/jcm10204661](https://doi.org/10.3390/jcm10204661)] [Medline: [34682779](https://pubmed.ncbi.nlm.nih.gov/34682779/)]
43. Ghomrawi HMK, Franco Ferrando N, Mandl LA, Do H, Noor N, Gonzalez Della Valle A. How often are patient and surgeon recovery expectations for total joint arthroplasty aligned? Results of a pilot study. *HSS J*. 2011;7(3):229-234. [FREE Full text] [doi: [10.1007/s11420-011-9203-6](https://doi.org/10.1007/s11420-011-9203-6)] [Medline: [23024618](https://pubmed.ncbi.nlm.nih.gov/23024618/)]
44. Klifto K, Klifto C, Slover J. Current concepts of shared decision making in orthopedic surgery. *Curr Rev Musculoskelet Med*. 2017;10(2):253-257. [FREE Full text] [doi: [10.1007/s12178-017-9409-4](https://doi.org/10.1007/s12178-017-9409-4)] [Medline: [28337730](https://pubmed.ncbi.nlm.nih.gov/28337730/)]
45. Henry S, Bell RA, Fenton JJ, Kravitz RL. Goals of chronic pain management: do patients and primary care physicians agree and does it matter? *Clin J Pain*. 2017;33(11):955-961. [FREE Full text] [doi: [10.1097/AJP.0000000000000488](https://doi.org/10.1097/AJP.0000000000000488)] [Medline: [28244944](https://pubmed.ncbi.nlm.nih.gov/28244944/)]
46. Pattabhiraman NK, Barnes SC, Singh HP. Agreement between patient-based and clinician-based assessment of the shoulder. *J Shoulder Elbow Surg*. 2018;27(3):e59-e67. [doi: [10.1016/j.jse.2017.09.008](https://doi.org/10.1016/j.jse.2017.09.008)] [Medline: [29128374](https://pubmed.ncbi.nlm.nih.gov/29128374/)]
47. Belayneh R, Lott A, Haglin J, Zuckerman J, Egol K. The role of patients' overall expectations of health on outcomes following proximal humerus fracture repair. *Orthop Traumatol Surg Res*. 2021;107(8):103043. [FREE Full text] [doi: [10.1016/j.otsr.2021.103043](https://doi.org/10.1016/j.otsr.2021.103043)] [Medline: [34389496](https://pubmed.ncbi.nlm.nih.gov/34389496/)]
48. Whitebird RR, Solberg LI, Ziegenfuss JY, Asche SE, Norton CK, Swiontkowski MF, et al. Personalized outcomes for hip and knee replacement: the patients point of view. *J Patient Rep Outcomes*. 2021;5(1):116. [FREE Full text] [doi: [10.1186/s41687-021-00393-z](https://doi.org/10.1186/s41687-021-00393-z)] [Medline: [34735662](https://pubmed.ncbi.nlm.nih.gov/34735662/)]
49. Wright AA, Hensley CP, Gilbertson J, Leland JM, Jackson S. Defining patient acceptable symptom state thresholds for commonly used patient reported outcomes measures in general orthopedic practice. *Man Ther*. 2015;20(6):814-819. [doi: [10.1016/j.math.2015.03.011](https://doi.org/10.1016/j.math.2015.03.011)] [Medline: [25843266](https://pubmed.ncbi.nlm.nih.gov/25843266/)]
50. Burke H, Sessums LL, Hoang A, Becher DA, Fontelo P, Liu F, et al. Electronic health records improve clinical note quality. *J Am Med Inform Assoc*. 2015;22(1):199-205. [FREE Full text] [doi: [10.1136/amiainl-2014-002726](https://doi.org/10.1136/amiainl-2014-002726)] [Medline: [25342178](https://pubmed.ncbi.nlm.nih.gov/25342178/)]
51. Yadav S, Kazanji N, Paudel S, Falatko J, Shoichet S, Maddens M, et al. Comparison of accuracy of physical examination findings in initial progress notes between paper charts and a newly implemented electronic health record. *J Am Med Inform Assoc*. 2017;24(1):140-144. [FREE Full text] [doi: [10.1093/jamia/ocw067](https://doi.org/10.1093/jamia/ocw067)] [Medline: [27357831](https://pubmed.ncbi.nlm.nih.gov/27357831/)]
52. Ebbers T, Kool RB, Smeele LE, Dirven R, den Besten CA, Karssemakers LHE, et al. The impact of structured and standardized documentation on documentation quality; a multicenter, retrospective study. *J Med Syst*. 2022;46(7):46. [FREE Full text] [doi: [10.1007/s10916-022-01837-9](https://doi.org/10.1007/s10916-022-01837-9)] [Medline: [35618978](https://pubmed.ncbi.nlm.nih.gov/35618978/)]

53. Hanson JL, Stephens MB, Pangaro LN, Gimbel RW. Quality of outpatient clinical notes: a stakeholder definition derived through qualitative research. *BMC Health Serv Res.* 2012;12:407. [FREE Full text] [doi: [10.1186/1472-6963-12-407](https://doi.org/10.1186/1472-6963-12-407)] [Medline: [23164470](https://pubmed.ncbi.nlm.nih.gov/23164470/)]
54. Jones JM, McPherson CJ, Zimmermann C, Rodin G, Le LW, Cohen SR. Assessing agreement between terminally ill cancer patients' reports of their quality of life and family caregiver and palliative care physician proxy ratings. *J Pain Symptom Manage.* 2011;42(3):354-365. [FREE Full text] [doi: [10.1016/j.jpainsymman.2010.11.018](https://doi.org/10.1016/j.jpainsymman.2010.11.018)] [Medline: [21454041](https://pubmed.ncbi.nlm.nih.gov/21454041/)]

## Abbreviations

**EHR:** electronic health record

**ICD10:** International Classification of Diseases, 10th Revision

**PES:** Patient Engagement Studio

**PHF:** proximal humerus fracture

**REDCap:** Research Electronic Data Capture

*Edited by T Leung; submitted 06.05.24; peer-reviewed by AJ Taxter, M Silberberg; comments to author 19.10.24; revised version received 14.11.24; accepted 06.12.24; published 23.01.25*

*Please cite as:*

*Floyd SB, Sutton JC, Okon M, McCarthy M, Fisher L, Judkins B, Reynolds ZC, Kennedy AB*

*Assessing Physician and Patient Agreement on Whether Patient Outcomes Captured in Clinical Progress Notes Reflect Treatment Success: Cross-Sectional Study*

*J Particip Med 2025;17:e60263*

URL: <https://jopm.jmir.org/2025/1/e60263>

doi: [10.2196/60263](https://doi.org/10.2196/60263)

PMID:

©Sarah B Floyd, Jordyn C Sutton, Marvin Okon, Mary McCarthy, Liza Fisher, Benjamin Judkins, Zachary Cole Reynolds, Ann Blair Kennedy. Originally published in *Journal of Participatory Medicine* (<https://jopm.jmir.org>), 23.01.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in *Journal of Participatory Medicine*, is properly cited. The complete bibliographic information, a link to the original publication on <https://jopm.jmir.org>, as well as this copyright and license information must be included.