Comparing Outcomes of Musculoskeletal Radiographs from In-Person and Telemedicine Primary Care Cohorts, April 2019–June 2021

John Petrilli, MD, FAAFP, Taylor Guth, BS, and Emily Coughlin, MPH

Purpose: Musculoskeletal conditions are a common reason for primary care visits, and they are being increasingly addressed at televisits. We therefore examined outcomes of musculoskeletal radiographs ordered at in-person and telemedicine primary care visits, which have implications for patient care and the economic impact of telemedicine.

Methods: We performed a retrospective cohort study of musculoskeletal radiograph orders placed April 1, 2019–March 31, 2021 at a major academic health system. Radiology reports were classified as normal or abnormal based on the radiologist’s impression. Findings were compared using c² tests.

Results: The main outcome was radiographic abnormalities. A secondary outcome was the effect of social determinants of health and medical comorbidities on telemedicine utilization. A total of 1580 radiographs were reviewed. Compared with televisits occurring after onset of the SARS-CoV-2-19 pandemic, radiographs ordered at in-person visits had higher odds of being abnormal (OR 2.51, 95% CI 1.33–4.75; \( P = .004 \)). When comparing radiographic outcomes at in-person visits before and after the pandemic’s onset, those ordered afterward had higher odds of being abnormal (OR 1.88, 95% CI 1.30–2.71; \( P < .001 \)). Social determinants of health and medical comorbidities were not associated with telemedicine utilization.

Conclusions: After the onset of the SARS-CoV-2-19 pandemic, radiographs ordered at in-person visits had higher odds of being abnormal compared with televisits. These findings indicate that prudence should be applied to ordering musculoskeletal radiographs in telemedicine encounters. (J Am Board Fam Med 2023;36:739–745.)

Keywords: Clinical Medicine, Delivery of Health Care, Family Medicine, Musculoskeletal Diseases, Orthopedics, Primary Health Care, Radiography, Retrospective Studies, Social Determinants of Health, Telemedicine

Introduction

Musculoskeletal conditions are a common reason for primary care encounters, with approximately 10 to 20% of the population presenting annually for a musculoskeletal complaint.¹,² However, the SARS-CoV-2-19 pandemic led to rapid changes in the delivery of primary care with more frequent utilization of telemedicine,³ and studies show that the content of televisits differs from that of in-person visits.⁴,⁵ Primary care clinicians are deciding which conditions are best managed via televisits, though data on actual outcomes is lacking.⁶–⁸ Clinicians are unsure if telemedicine is an appropriate medium to evaluate musculoskeletal complaints, largely due to the difficulty performing a complex physical examination.⁹,¹⁰ Efforts are being made to train primary care physicians to perform virtual musculoskeletal exams¹¹ and to evaluate the effectiveness of virtual musculoskeletal exams.¹² However, to our knowledge there are no
studies comparing outcomes for musculoskeletal radiographs ordered at virtual visits to in-person visits. Primary care physicians order radiographs at up to 25% of encounters for musculoskeletal conditions, so the implications for patient care and costs to the health system and society are not insignificant.

Therefore, we examined the results of radiographs ordered at in-person and telemedicine visits before and after onset of the pandemic. People with low English proficiency, age 65 or older, Black race, Latinx ethnicity, female sex, lack of prior telecommunications use, lower household income, and enrollment in Medicaid have been shown to use telemedicine less than others. Therefore, as a secondary outcome we explored the association between social determinants of health and medical comorbidities with telemedicine utilization. We hypothesized that radiographs ordered at televisits would have fewer radiographic abnormalities than those ordered at in-person visits. In addition, we hypothesized that patients using telemedicine during the COVID-19 pandemic would be younger and English-speaking.

Methods
Study Design
The study utilized a retrospective cohort design. We queried the electronic health record from April 1, 2019 to March 31, 2021 at a large academic health system. Adult primary care patients were included if they had a visit containing ICD-10 codes beginning with M or S (musculoskeletal codes) for which a radiograph was ordered. The university’s IRB approved the study.

Data Collection
Patient characteristics included medical record number, race, ethnicity, language, gender, age, payor group, median household income based on postal code, and comorbid ICD-10 codes for calculating an age-adjusted Charleson comorbidity index (CCI). Characteristics of orders included ICD-10 code, anatomic area of the radiograph, date of the order, encounter type (telemedicine or in-person), and date radiograph was completed. Our institution’s policy was to use video televisits and convert to telephone-only if the video failed. We were unable to differentiate video from telephone televisits.

One of the authors personally reviewed the radiograph reports and coded the findings as normal or abnormal based on the radiologist’s impression. If there was uncertainty regarding the impression, we conferred and came to consensus. The primary finding was also recorded, as documented by the radiologist’s impression on the radiology report. When applicable, the primary finding’s severity was included as stated by the radiology report (eg, moderate osteoarthritis). The authors categorized abnormal findings into groupings including spinal spondylosis, fracture, mild osteoarthritis, moderate or advanced osteoarthritis, other diagnosis, and missing result.

Data Analysis
Patient characteristics were analyzed only for unique patients, because some patients underwent multiple radiographs. Categorical data were summarized using frequencies and proportions, whereas continuous data were summarized as mean ± standard deviation. Chi-square tests were used to compare categorical variables, whereas due to non-normal distribution, nonparametric Mann-Whitney U test was used to compare continuous variables of age, household income, and age-adjusted CCI.

Radiographs were analyzed per-order and organized into 3 categories: those ordered at in-person visits before April 1, 2020, in-person visits on or after April 1, 2020, and televisits on or after April 1, 2020. We chose April 1, 2020, as the cutoff, because that is when our institution began using telemedicine. Radiographs were analyzed per-order instead of per-patient, so as to avoid introduction of selection bias for patients who underwent multiple radiographs. Orders with missing results were excluded. Proportions of abnormal findings were compared using Chi-square tests. Results were reported as Pearson’s c² p-values and Mantel Haenszel odds ratios (OR) with 95% confidence intervals (95% CI). OR were unadjusted as none of the patient factors were statistically significant. All p-values less than α = 0.05 were considered statistically significant. Analysis was conducted using SPSS Version 27.

Results
A total of 1580 radiograph orders were analyzed for 995 unique patients. Patients were mostly female (n = 623, 63%), white (n = 668, 69%), and non-Latinx (n = 831, 86%). The average age was 62.9 ± 15.2 years. Table 1 displays patient characteristics and their
relationship to telemedicine utilization. Radiographs done before April 1, 2020 were all ordered at in-person visits (n = 812, 51%). Radiographs done after April 1, 2020 (n = 768) were ordered at either in-person (n = 636, 40%) or telemedicine visits (n = 132, 8%). Among patients who visited after April 1, 2020 (n = 507), 89 (18%) utilized telemedicine. Patients who are non-Latinx were more likely to use telemedicine (P = 0.046), although the odds ratio was not statistically significant (OR 0.444, 95% CI 0.196–1.01).

We conducted a qualitative post hoc review of results for patients who underwent multiple radiographs. Multiple radiographs were ordered on 371 of the patients over the course of the 2-year span. Of these, 72 pairs of radiographs either overlapped anatomic sites or were of the same site at separate points in time. Sixty-seven of the pairs were ordered at in-person appointments, and 49 of these pairs had concordant abnormal results. The abnormal results were relatively equally distributed before and after the pandemic’s onset (22 pairs each), and 5 of the pairs were separated by the pandemic’s onset.

Table 2 compares radiograph findings between the 3 visit-types. Results were available for 1075 orders (68%). Radiographs from in-person visits after the pandemic’s onset showed a significantly increased odds of abnormal findings compared with televisits (OR 2.51, 95% CI 1.33–4.75; P = 0.004). Radiographs ordered at in-person visits after the pandemic’s onset showed a significantly higher odds of abnormal findings compared with those ordered before the pandemic’s onset (OR 1.88, 95% CI 1.30–2.71; P < 0.001). Table 3 shows the proportion of normal and abnormal radiograph diagnoses among the groups, including a breakdown of abnormal findings. “Fracture” was by far the most common severe finding. There were 2 radiographs with erosions and 1 radiograph with evidence of osteomyelitis, which were included in “other diagnoses” due to their low frequency.
Discussion

The literature has repeatedly called for studies evaluating actual outcomes of televisits.6–8 Yet in 2022, the authors of a systematic review of orthopedic televisits still concluded that “further studies are needed to determine whether telemedicine consultations from the patient’s own home are truly reliable and achievable.”18 To our knowledge, this study is the first to explore outcomes of radiographs ordered at televisits.

The results indicate that primary care physicians may overutilize musculoskeletal radiographs at televisits. A possible explanation for these results is that, during the height of the pandemic, patients with minor symptoms avoided in-person visits, whereas patients with more severe conditions sought in-person care. The idea of patient self-selection accords with a survey of primary care patients who delayed care for orthopedic concerns that arose during the pandemic.19 A study of Medicare patients found that up to 45% of them avoided medical facilities during the pandemic for fear of contracting the virus.20 Delaying care may have inflated radiographic abnormalities for in-person visits that occurred after the height of the pandemic. Ongoing surveillance is needed to see if this trend persists.

Alternatively, physicians may have overutilized radiographs at televisits to compensate for the lack of a tactile physical examination. Unfortunately, the literature on imaging outcomes from virtual visits is sparse. A study looking specifically at patients presenting for low back pain found that less imaging was ordered for patients presenting virtually compared with in-person.21 However, national guidelines recommend against routine imaging for low back pain,22 and the

| Table 2. Comparisons between Outcomes for Radiographs Ordered at In-Person and Telemedicine Visits, Pre- and Post-Pandemic Onset |
|---|---|---|---|
| Comparisons^a | Normal Findings^b | Abnormal Findings^b | OR (95% CI) | p-Value |
| 1. Post-onset, in-person versus telemedicine |
| Post-onset, in-person | 47 (10.9) | 384 (89.1) | 2.51 (1.33–4.75) | 0.004 |
| Post-onset, telemedicine | 16 (23.5) | 52 (76.5) | | |
| 2. In-person, post-versus pre-onset |
| Post-onset, in-person | 47 (10.9) | 384 (89.1) | 1.88 (1.30–2.71) | <0.001 |
| Pre-onset, in-person | 107 (18.7) | 466 (81.3) | | |

Abbreviations: OR, Odds ratio; CI, Confidence interval.
^aDoes not include missing data.
^bData presented as N (%).

Table 3. Outcomes of Musculoskeletal Radiographs by Visit Type, and Breakdown of the Primary Finding of Abnormal Radiographs

<table>
<thead>
<tr>
<th></th>
<th>Before Onset, In-person^a</th>
<th>After Onset, In-person^a</th>
<th>After Onset, Telemedicine^a</th>
</tr>
</thead>
<tbody>
<tr>
<td>Missing</td>
<td>239</td>
<td>205</td>
<td>64</td>
</tr>
<tr>
<td>Normal findings</td>
<td>107 (18.7)</td>
<td>47 (10.9)</td>
<td>16 (23.5)</td>
</tr>
<tr>
<td>Abnormal findings</td>
<td>466 (81.3)</td>
<td>384 (89.1)</td>
<td>52 (76.5)</td>
</tr>
<tr>
<td>Primary abnormality:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spinal spondylosis</td>
<td>157 (27.4)</td>
<td>95 (22.0)</td>
<td>17 (25.0)</td>
</tr>
<tr>
<td>Fracture</td>
<td>20 (3.5)</td>
<td>29 (6.7)</td>
<td>2 (2.9)</td>
</tr>
<tr>
<td>Mild osteoarthritis</td>
<td>140 (24.4)</td>
<td>119 (27.6)</td>
<td>16 (23.5)</td>
</tr>
<tr>
<td>Moderate/advanced osteoarthritis</td>
<td>88 (15.4)</td>
<td>83 (19.3)</td>
<td>12 (17.6)</td>
</tr>
<tr>
<td>Other diagnosis^b</td>
<td>61 (10.6)</td>
<td>58 (13.5)</td>
<td>5 (7.4)</td>
</tr>
</tbody>
</table>

^aData presented as N(%). Percentages are calculated excluding missing results.
^bOther diagnoses included calcific tendinitis, enthesopathy, erosions (2 radiographs), joint effusion, osteomyelitis (1 radiograph), osteopenia, postsurgical changes, and soft tissue changes.
study did not explicitly look at imaging results. A study of virtual and in-person encounters for abdominal complaints in the ER found that, compared with the in-person evaluation, clinicians performing the virtual evaluation ordered less imaging. Records from the ER televisits documented greater diagnostic uncertainty, and the authors hypothesized that concerns about cost and imaging utilization may have led the virtual clinicians to order less imaging. Clinical uncertainty may also explain the results of our study. Clinicians may have ordered imaging at televisits to compensate for the lack of a reliable physical examination. A virtual musculoskeletal examination has been shown to provide good accuracy for elements like range of motion, but special tests are more difficult to perform. Multiple studies document the challenges that physicians and patients face when engaging in a virtual musculoskeletal examination, including the potential need for caregiver assistance, appropriateness of the patient’s environment, poor visualization, low inter-rater reliability, and technical issues such as video lag. Clinicians should not underestimate the harms of an incomplete or inaccurate physical examination, though a growing familiarity with the virtual musculoskeletal examination might impact future ordering patterns for radiographs. Forthcoming studies could evaluate which components of the in-person visit (e.g., history and/or physical examination) best correlate with the need for musculoskeletal imaging.

In an age mindful of cost and resource utilization, we should also consider the economic impact of primary care televisits for musculoskeletal conditions. Prior studies exploring health systems and societal costs of orthopedic televisits did not specifically consider the cost of imaging; they either ignored imaging or assumed costs were similar for televisits. Our study’s findings challenge this assumption. Telemedicine does not routinely reduce the cost of health care, largely due to the costs associated with technology upkeep and maintenance. It has the potential to reduce health care costs if it mitigates expensive procedures and specialist referral, but our study indicates that televisits may be less than ideal for musculoskeletal conditions. Unnecessary radiography would increase health care costs (in 2022 CMS reimbursed $34.26 for CPT 73030, multi-view shoulder radiograph) and further mitigate any savings to the health system gained by the televisit. Furthermore, telemedicine’s societal benefits are largely attributable to the decreased costs of travel and missed work. If telemedicine utilization is associated with unnecessary imaging orders, however, the need for travel to an imaging facility and time off of work would moderate the savings to society. This study found no significant association between social determinants of health or comorbidities and telemedicine utilization, which contrasts with much of the existing literature. Local factors and the population studied likely account for this. We did see trends toward significance with age and ethnicity.

Our study has several limitations. As telemedicine becomes a regular medium for ambulatory visits, care must be taken to contextualize these results, which were obtained during the height of the COVID-19 pandemic. A major limitation of this study is that imaging outcomes were not compared with a patient’s documented symptoms in the medical record. Radiographic abnormalities may therefore be incidental findings. However, any error introduced from this would likely be nondifferential, because incidental findings would be present in all groups. Another limitation of the study involves the per-order analysis approach. A per-patient approach could have been used, but it would be more likely to introduce selection bias, because only 1 radiograph per patient could be chosen. Qualitative review of the results found that, of the patients who obtained multiple radiographs of a similar anatomic site, the proportion of concordant abnormal results were similar before and after the pandemic’s onset. Lastly, as a retrospective cohort study, we could not control the size of the cohorts. The number of radiographs from televisits were relatively small compared with those ordered at in-person visits. Nevertheless, retrospective data such as this is valuable, because there is so little evidence of telemedicine’s actual outcomes. In addition, the unpredictability of patients’ concerns makes a prospective primary care cohort difficult to arrange, and randomizing musculoskeletal concerns to telemedicine or in-person visits would be difficult to do in the primary care setting.

In summary, radiographs ordered at in-person primary care visits after the pandemic’s onset had higher odds of being abnormal compared with televisits. These findings indicate that prudence should
be applied to ordering musculoskeletal radiographs in telemedicine encounters.

The authors would like to acknowledge Dr. Rahul Mhaskar for his contributions to the statistical analysis.

To see this article online, please go to: http://jabfm.org/content/36/5/739.full.

References


