Clinical Study

ASJ Asian Spine J. March 4, 2025 [Epub ahead of print]: 1-6
https://doi.org/10.31616/asj.2024.0344
PISSN 1976-1902 • eISSN 1976-7846

To Zoom or not to Zoom? A retrospective comparative study to assess the validity of video versus face-to-face consultations and examinations for diagnosing lower back pain in Wales

Megan Havard[1,2], Ronan McKeogh[3], Abi Goodier[3], Gemma Johns[2], Sashin Ahuja[4]

[1]Southmead Hospital, North Bristol NHS Trust, Bristol, UK
[2]TEC Cymru, Cardiff, UK
[3]Cardiff University School of Medicine, Cardiff, UK
[4]Welsh Centre for Spinal Surgery & Trauma, Cardiff, UK

Received Aug 29, 2024; **Revised** Oct 27, 2024; **Accepted** Nov 20, 2024
Corresponding author: Megan Havard
Southmead Hospital, Southmead Road, Westbury-on-Trym, Bristol, BS10 5NB, UK
E-mail: megan.havard@nbt.nhs.uk

Study Design: A retrospective comparative study.
Purpose: To evaluate the accuracy of examination findings and diagnoses established after initial spinal consultations for low back pain conducted virtually in comparison with face-to-face (F2F) consultations.
Overview of Literature: The COVID-19 pandemic required a shift from F2F to virtual consultations in healthcare, with documented benefits such as reduced costs, CO_2 emissions, and time. However, the diagnostic accuracy of telemedicine for conditions requiring physical examinations, such as low back pain, remains underexplored. Although studies have highlighted the feasibility and high satisfaction of virtual spinal assessments, they have not thoroughly investigated their diagnostic accuracy.
Methods: This study included 154 new patients with degenerative lumbar spine problems who were assessed via virtual consultations (VCs) (n=77) or F2F (n=77) by a single orthopedic spinal surgeon between May 2020 and January 2021. Patients were matched by age and sex, and examinations followed the "telemedicine musculoskeletal examination" protocol by the Mayo Clinic. Diagnostic changes from initial to definitive diagnoses were recorded and compared between the two groups.
Results: The diagnostic accuracy of VCs was comparable with that of F2F examinations, with no significant difference in the rate of diagnostic changes between the groups (p=0.814 for any change; p=0.405 for more significant changes). Motor deficits were the only examination component with significant false-positive rates in the VC group (p=0.023).
Conclusions: The study findings indicate that initial spinal VCs and examinations are effective, safe, and beneficial in the evaluation of low back pain, providing the same diagnostic accuracy as initial F2F consultations. All erroneous motor deficit findings were false-positives, which means that no serious pathology was missed during the initial VC evaluation. Future research should focus on refining virtual examination techniques, particularly for assessing motor deficits.

Keywords: Spine; Lower back pain; Examination; Diagnosis; Telemedicine

Introduction

The COVID-19 (coronavirus disease 2019) pandemic has precipitated a rapid change in the delivery of healthcare services: shifting away from face-to-face (F2F) consultations toward video consultations (VCs) [1-3]. Initially, this was done out of necessity to protect patients and clinicians and adhere to lockdown regulations; however, VCs have presented numerous additional benefits, such as lowering patients' costs, carbon footprint, and time [4,5]. Nevertheless, VCs have raised questions regarding safety and diagnostic accuracy [6,7], particularly the inability to perform hands-on examinations [8]. Over an online platform, the clinician cannot physically interact with the patient; therefore, the examination technique must be modified. In 2020, a systematic review assessed the validity of performing spinal assessments through telemedicine, yielding three studies predating the pandemic. At this time, no guidance has been established on performing a virtual examination, and they found a literature gap and recommended further research to validate virtual physical examinations for lower back pain (LBP) [9]. Based on experiences during the pandemic, various guidelines have now been published, recommending approaches for conducting virtual spinal examinations [10-15]. In addition, studies have used patient and clinician surveys to demonstrate the feasibility and high satisfaction of these virtual spinal assessments [16-19]; however, few have considered the diagnostic accuracy of VCs.

This primary aim of this study was to evaluate whether the diagnostic accuracy of virtual spinal examinations differed from those conducted during F2F clinics in patients presenting with LBP. The secondary aim was to quantify the benefit of reduced carbon footprint from traveling.

Materials and Methods

The study received ethical approval via Technology-Enabled Care (TEC) Cymru for all service evaluations conducted in association with the use of the NHS Wales VC Service (reference no., SA/1114/20).

Retrospective data were collected from consecutive new patients presenting with lumbar spine problems were seen through VCs with a single orthopedic spinal surgeon practicing in Wales between May 2020 and January 2021. The data included demographic details, examination findings, clinical diagnosis based on the initial VC, imaging results, and subsequent F2F examination findings. Patients who had an initial VC without F2F follow-up or subsequent imaging were excluded. The patient group was then matched by age and sex to patients within the same database (i.e., assessed by the same surgeon in the same outpatient clinic) seen either before or after the pandemic for an initial F2F appointment (2006–2023). This resulted in a total of 154 patients, with 77 in each group. The examinations took place on Zoom (Zoom Video Communications Inc., San Jose, CA, USA), following the structure of the "telemedicine musculoskeletal examination" (TME) protocol by the Mayo Clinic as a template [11], ensuring consistency. Although the protocol covers various spinal tests, only the components that we deemed the most user-friendly were replicated, namely, the range of motion (ROM), muscle strength, and straight-leg raise (SLR). The TME protocol does not include the femoral nerve stretch test (FNST); the method we used for this required patients to stand with a balance aid, bend their knee to 90°, and then extend their hip. Anterior thigh pain was considered a positive finding.

In both groups, the initial clinical diagnosis after the initial appointment by VC or F2F was compared, and a definitive diagnosis was made after subsequent imaging and further examinations. The individual components of the spinal examination were also considered, and any differences between the virtual examination and subsequent F2F examination findings were noted. The components compared were ROM, motor deficit, SLR, and FNST. The results of the SLR and FNST test were categorized as a binary outcome of "positive" or "negative" rather than recording the exact angles, and a sensory examination was not included because it was considered too subjective. The initial and subsequent examination findings were also compared in the F2F group. Pearson chi-square tests were used to compare the number of changes in each group for both diagnoses and examination findings.

To achieve the secondary outcome, using the shortest mileage route of Google Maps, the home postcodes of the VC group were used to calculate the distance they would have had to travel to the hospital if they had commuted.

Results

Of the 154 patients included, 51.3% were female and 48.7% were male, with an equal distribution in both groups because of sex matching. In each group, the ages of patients were matched within 6 months, resulting in an average age of 49 and 50 in the F2F and VC groups, respectively, at the date of their first appointment. The

most common definitive diagnoses in both groups were radiculopathies and nonspecific LBP, and the respective numbers of patients are shown in Table 1.

Diagnostic accuracy

In the VC group, the initial clinical diagnoses of 10 patients changed after having an F2F examination and, in some cases, imaging. These included a change in the diagnosis of radiculopathy from a spinal level to its adjacent spinal level, which was considered a minor change, for example, a provisional diagnosis of L5 radiculopathy became S1 radiculopathy. These changes are common and generally acceptable in clinical examination and imaging because of differences in human anatomy, overlap in nerve distributions, and subjectivity in the descriptions of symptoms. Of the 10 patients, two had a more significant change: initially, they were diagnosed with nonspecific LBP (i.e., with no neural issues); however, subsequent magnetic resonance imaging showed nerve compression. In the F2F group, 11 diagnoses changed after imaging was performed, with seven of these being a minor change (between adjacent spinal levels) and four being more significant, i.e., from pain with a neuropathic origin to pain with vascular or mechanical origin. Table 2 shows the number of diagnostic changes in each group. When considering those with any change between clinical and definitive diagnosis, the difference between the two groups was not significant (p=0.814); similarly, no difference was noted when analyzing only those with more clinically significant changes (p=0.405).

Table 1. Number of patients with each definitive diagnosis

Definitive diagnosis	Face-to-face (n=77)	Online (n=77)	Total (n=144)
Radiculopathy	39	41	80 (55.6)
Non-specific lower back pain	37	32	69 (47.9)
Scoliosis	0	2	2 (1.4)
Pectus	0	1	1 (0.7)
Lipoma	0	1	1 (0.7)
Vascular claudication	1	0	1 (0.7)

Values are presented as number or number (%).

Table 2. Number of patients with a change in initial to definitive diagnosis in each group

	No change	Minor change	Significant change	Total
Video consultations	67 (87)	8 (10)	2 (3)	77 (100)
Face-to-face	66 (86)	7 (9)	4 (5)	77 (100)

Values are presented as number (%).

Examination accuracy

With regard to physical examination findings, all 77 patients in the F2F group had completed all examination components, and no changes in the findings were noted between the initial and subsequent examinations. Among the 77 patients in the VC group, some aspects of the examination either had not been completed or data were missing; accordingly, they were excluded from comparisons for that particular component. Table 3 shows the number of patients in the VC group with changes in each examination component when subsequently examined in person. No variations were observed in the ROM assessments. Of the 68 patients who underwent the FNST online, 1 (1.5%) tested positive in the VC but tested negative in the F2F examination. When compared with the 100% accuracy in the F2F group, the difference was not significant (p=0.286). Of the 74 patients who performed the SLR virtually, 2 (2.7%) exhibited different findings in the subsequent F2F examinations. However, when compared with the F2F group, the difference was also not significant (p=0.146).

There was, however, a clinically significant false-positive rate of motor deficits noted during VCs. That is, five of the 77 patients demonstrated motor deficits during the VC; however, such findings were not observed in subsequent F2F examinations, equating to a 6.5% change when compared with 0% for the F2F group, which indicated a significant difference (p=0.023).

Discussion

To our knowledge, this is the first study that evaluates the accuracy of the examination techniques described in the TME of the Mayo Clinic (with the addition of our FNST technique) when applied to the initial LBP assessment by VCs in comparison with a control group who underwent initial F2F consultations. Our results shed light on the feasibility of virtual spinal consultations, which yielded an initial diagnostic accuracy equal

Table 3. Number of patients who had different findings on initial video consultations compared to subsequent face-to-face examination

	No change	Change	Total
Range of motion	77 (100.0)	0	77 (100.0)
Motor deficits	**72 (93.5)**	**5 (6.5)**	77 (100.0)
Femoral nerve stretch test	67 (98.5)	1 (1.5)	68 (100.0)
Straight-leg raise	72 (97.3)	2 (2.7)	74 (100.0)

Values are presented as number (%). Statistically significant results are marked in bold.

to that of the initial F2F evaluation.

The findings of this study align with the recent research. For example, in 2021, Crawford et al. [20] reported that procedural plans remained unchanged after F2F evaluation in 87% of 87 patients seen by interventional spine specialists through VCs, consistent with our 87% unchanged diagnosis rate. A similar study published in 2022 unveiled that the diagnosis was unchanged in 91% of 65 patients with spinal problems who had a VC appointment and a subsequent F2F visit, and the surgical plan was unchanged in 80% of the patients [21]. Notably, these studies did not include a comparison with a F2F group, preventing the assessment of the significance of this accuracy. Another recent study, which compared the initial surgical plans made by spinal surgeons in F2F clinics and virtually, did not find a significant difference in the accuracy of surgical plans [22]. Although the present study did not specifically focus on surgical plans, diagnostic accuracy is closely related to it, and the existing focus on surgical plans demonstrates the relevance of our distinct research focus.

With regard to the reliability of physical examination findings determined at VCs, in 2021, Jansen et al. [23] reported on virtual spinal examinations performed on 43 patients who subsequently underwent F2F examinations for comparison of findings. They found that virtual ROM assessments had "good" agreement with F2F examinations, which corresponds to our findings that 100% of the findings in virtual ROM assessments were unchanged when examined in person. In contrast to our findings, they found only "moderate" agreement in the SLR test results but "good" to "very good" agreement in the neurophysiological component, which included power. They did not perform the FNST. The lack of established guidelines for virtual spinal examinations when this study was conducted may account for the discrepancies in our results.

In this study, only the motor deficit assessment demonstrated a significant difference between the VC and F2F groups. Motor deficits were examined by a double leg squat, which may be a more difficult exercise than performing isolated movements against the examiner's resistance in a standard neurological examination, potentially leading to false-positive findings. This was not the case for methods by Jansen et al. [23] of assessing power. Although motor deficits were overidentified in our initial VC, all erroneous findings were false-positives. This reassurance indicates that no motor deficits and therefore no serious pathologies were missed during the initial VCs. Furthermore, none of the five patients with a false-positive motor deficits in the initial VC had a significant change in diagnosis at subsequent F2F evaluations. However, two patients had minor changes in radiculopathy levels, i.e., from S1 to L5 and from L5 to both L4 and L5. This weakness is not significant enough to deter offering initial VCs for LBP assessment. However, the field would benefit from establishing more reliable methods for assessing power during VCs. All other examination components were not significantly different between VC and F2F examinations, and this limitation of VCs did not affect the most important outcome assessed in this study: the accuracy of the initial diagnosis made during VCs. An international survey in May 2020 found that 38.6% of spine surgeons identified the ability to perform physical examinations as a major obstacle to telemedicine [24]. This rate is reassuringly quite low and hopefully will be even lower now with more experience and guidance, such as this study. However, considering the modality of service provision, not only the accuracy of the examination or diagnosis is important but also the logistical challenges surrounding VCs, such as internet access, patients' digital literacy, confidence, access to a device with a camera in a suitable room to complete the examination, and patient preference [25,26]. F2F should always be an option if the patient wishes. Patients with comorbidities, such as those who are deaf or blind or have learning disabilities, are also considered. During the pandemic, visual aids, diagrams, and brochures could be used to aid patients' understanding of what is being asked of them in a spinal examination [3]. However, if a patient requires this much additional assistance, it is often easier to examine them in person.

Regarding our secondary objective, each VC saved 3.23 kg of carbon dioxide (CO_2) based on an average patient home-to-hospital distance of 14.6 miles, and data from the Department of Transport indicating that, on average, a car in the United Kingdom emits 221.4 g of CO_2 per mile [27]. Although this is specific to location and context, it aligns with the finding of a systematic review that a VC can save between 0.70 and 372 kg of CO_2 [28]. The lower bound for the carbon footprint of a 1-hour Zoom call is 150 g of CO_2, which is negligible compared with the amount saved [29]. Thus, a widespread move to using video platforms could help significantly reduce greenhouse gas emissions, as has been acknowledged in the literature [30]. Moreover, whether through private or public transportation, the average distance of 14.1 miles from the hospital necessitates commuting time and associated costs, along with potential parking charges. This becomes more signifi-

cant for patients with multiple appointments, increasing both time and financial burdens.

Some limitations of this study should be considered. As a single-surgeon series, the results are less generalizable; nonetheless, intraobserver variability is reduced by having the same clinician perform all the examinations. In the VC group, certain examination components were not always completed due to patients' difficulties in following the assessments, meaning sample sizes for these components were reduced. Our comparative results only included those who had completed each test. With a 100% completion rate of the examinations in the F2F group, this presents a limitation of telemedicine itself. In addition, despite being a fairly routine test in F2F assessments, deep tendon reflexes were not examined virtually because we felt it may be challenging for patients to self-elicit reflexes as suggested in the TME. The lack of a suitable method to assess reflexes also demonstrates a limitation of VCs.

Conclusions

This study contributes to the growing body of evidence supporting the feasibility of virtual spinal consultations. The findings provide reassurance to patients and spinal surgeons undertaking VCs by demonstrating the accuracy of VCs for initial assessment and diagnosis. The benefits of VCs, such as time saving, convenience, cost reduction, and lower carbon footprint, are not necessarily compromised by an inferior evaluation. Although VCs present a safe and viable alternative to F2F assessments for LBP in the outpatient setting, it is not an entirely equivalent examination. Not all patients in the VC group had completed every aspect of the physical examination, and deep tendon reflexes were not tested. Moreover, there were false-positive cases in the motor deficit assessments. Therefore, more studies are necessary to explore the effectiveness and accuracy of VCs of lower limb power, and caution is advised when integrating VCs into clinical practice.

Key Points

- No significant difference was found in the diagnostic accuracy of virtual consultations for the initial assessment of low back pain compared with that of face-to-face consultations.
- Although the telemedicine musculoskeletal examination protocol by the Mayo Clinic may produce some false-positive motor deficits, it is a valid guide for performing spinal examinations.
- To virtually perform the femoral nerve stretch test, patients must stand with a balance aid, bend their knees to 90°, and then extend their hips. Anterior thigh pain is a positive finding. Compared with the conventional in-person method, the proposed virtual method did not show a significant difference in accuracy.

Conflict of Interest

No potential conflict of interest relevant to this article was reported.

ORCID

Megan Havard: https://orcid.org/0000-0001-9808-0528;
Ronan McKeogh: https://orcid.org/0009-0005-7237-5041;
Abi Goodier: https://orcid.org/0009-0000-4031-4803;
Gemma Johns: https://orcid.org/0000-0002-3983-362X;
Sashin Ahuja: https://orcid.org/0000-0001-8676-6100

Author Contributions

Conceptualization: SA. Data curation: MH, AG. Investigation: MH, AG. Formal analysis: MH. Visualization: RM. Validation: GJ. Writing–original draft: MH. Writing–review and editing: RM, GJ. Project administration: SA. Supervision: SA. Final approval of the manuscript: all authors.

References

1. Richardson E, Aissat D, Williams GA, Fahy N. Keeping what works: remote consultations during the COVID-19 pandemic. Eurohealth 2020;26:73-6.
2. Bokolo Anthony Jnr. Use of telemedicine and virtual care for remote treatment in response to COVID-19 pandemic. J Med Syst 2020;44:132.
3. Liu KG, Tan WL, Yip WL, Tan JH, Wong HK. Making a traditional spine surgery clinic telemedicine-ready in the

"new normal" of coronavirus disease 2019. Asian Spine J 2021;15:164-71.
4. Havard M. Welsh Public Opinion Survey 2023: what the data is telling us so far from the perspective of a medical student working with TEC Cymru. Cardiff: TEC Cymru, Welsh Government; 2023.
5. Rutherford E, Noray R, O hEarrain C, et al. Potential benefits and drawbacks of virtual clinics in general surgery: pilot cross-sectional questionnaire study. JMIR Perioper Med 2020;3:e12491.
6. Lee G, Clough OT, Hayter E, et al. Embracing virtual outpatient clinics in the era of COVID-19. Bone Jt Open 2021;2:301-4.
7. Greenhalgh T, Vijayaraghavan S, Wherton J, et al. Virtual online consultations: advantages and limitations (VOCAL) study. BMJ Open 2016;6:e009388.
8. Mazarakis NK, Koutsarnakis C, Komaitis S, Drosos E, Demetriades AK. Reflections on the future of telemedicine and virtual spinal clinics in the post COVID-19 era. Brain Spine 2022;2:100930.
9. Piche J, Butt BB, Ahmady A, Patel R, Aleem I. Physical examination of the spine using telemedicine: a systematic review. Global Spine J 2021;11:1142-7.
10. Iyer S, Shafi K, Lovecchio F, et al. The spine physical examination using telemedicine: strategies and best practices. Global Spine J 2022;12:8-14.
11. Laskowski ER, Johnson SE, Shelerud RA, et al. The telemedicine musculoskeletal examination. Mayo Clin Proc 2020;95:1715-31.
12. Yoon JW, Welch RL, Alamin T, et al. Remote virtual spinal evaluation in the era of COVID-19. Int J Spine Surg 2020;14:433-40.
13. Satin AM, Lieberman IH. The virtual spine examination: telemedicine in the era of COVID-19 and beyond. Global Spine J 2021;11:966-74.
14. Sardar ZM, Coury JR, Luzzi AJ, Weidenbaum M, Riew KD. The telehealth spine physical examination: a practical approach learned during the COVID-19 pandemic. World Neurosurg 2021;154:e61-71.
15. Haddad AF, Burke JF, Mummaneni PV, et al. Telemedicine in neurosurgery: standardizing the spinal physical examination using a modified Delphi method. Neurospine 2021;18:292-302.
16. Hobson S, Aleem IS, Bice MJ, et al. A multicenter evaluation of the feasibility, patient/provider satisfaction, and value of virtual spine consultation during the COVID-19 pandemic. World Neurosurg 2021;154:e781-9.
17. Craig M, Chopra A, Lasry O, et al. Telehealth for outpatient spine consultation: what do the patients think? Interdiscip Neurosurg 2022;28:101462.
18. Viswanathan VK, Shetty AP, Murugan C, Kavishkar RA, Kanna RM, Rajasekaran SB. The perspectives of patients and spine surgeons on the effectiveness and reliability of telemedicine consultations for spine ailments: a glimpse at the Indian Scenario. Indian Spine J 2023;6:82-8.
19. Wessell NM, Kleck C, Ou-Yang D, Goldstein CL, Burger EL, Patel VV. P41: The University of Colorado virtual physical exam for spine surgeons: a preliminary study. Spine J 2021;21(9 Supplement):S160.
20. Crawford AM, Lightsey HM, Xiong GX, et al. Interventional procedure plans generated by telemedicine visits in spine patients are rarely changed after in-person evaluation. Reg Anesth Pain Med 2021;46:478-81.
21. Bovonratwet P, Song J, Kim YE, et al. Telemedicine visits can generate highly accurate diagnoses and surgical plans for spine patients. Spine (Phila Pa 1976) 2022;47:1194-202.
22. Ye IB, Thomson AE, Donahue J, et al. Similar accuracy of surgical plans after initial in-person and telemedicine evaluation of spine patients. World Neurosurg 2022;164:e1043-8.
23. Jansen T, Gathen M, Touet A, et al. Spine examination during COVID-19 pandemic via video consultation. Z Orthop Unfall 2021;159:193-201.
24. Riew GJ, Lovecchio F, Samartzis D, et al. Spine surgeon perceptions of the challenges and benefits of telemedicine: an international study. Eur Spine J 2021;30:2124-32.
25. Donaghy E, Atherton H, Hammersley V, et al. Acceptability, benefits, and challenges of video consulting: a qualitative study in primary care. Br J Gen Pract 2019;69:e586-94.
26. Haimi M. The tragic paradoxical effect of telemedicine on healthcare disparities: a time for redemption: a narrative review. BMC Med Inform Decis Mak 2023;23:95.
27. Yurday E. Average CO2 emissions per car in the UK [Internet]. London: NimbleFins; 2024 [cited 2024 May 12]. Available from: https://www.nimblefins.co.uk/average-co2-emissions-car-uk
28. Purohit A, Smith J, Hibble A. Does telemedicine reduce the carbon footprint of healthcare?: a systematic review. Future Healthc J 2021;8:e85-91.
29. Obringer R, Rachunok B, Maia-Silva D, Arbabzadeh M, Nateghi R, Madani K. The overlooked environmental footprint of increasing Internet use. Resour Conserv Recycl 2021;167:105389.
30. Pickard Strange M, Booth A, Akiki M, Wieringa S, Shaw SE. The role of virtual consulting in developing environmentally sustainable health care: systematic literature review. J Med Internet Res 2023;25:e44823.