Validity of movement smoothness metrics for upper limb reaching movements in people with moderate to severe subacute stroke

Gwenaël CORNEC (✉ gwenael.cornec@chu-brest.fr)
Department of Physical and Rehabilitation Medicine, Brest University Hospital

Mathieu LEMPEREUR
Department of Physical and Rehabilitation Medicine, Brest University Hospital

Johanne MENSAH-GOURMEL
Department of Physical and Rehabilitation Medicine, Brest University Hospital

Johanna ROBERTSON
Physical and Rehabilitation Medicine, Raymond Poincaré Hospital, AP-HP, Paris-Saclay University, INSERM 1179

Ludovic MIRAMAND
Laboratory of Medical Information Processing (LaTIM), UMR 1101 INSERM, University of Brest

Beatrice MEDEE
Department of Physical and Rehabilitation Medicine, Brest University Hospital

Soline BELLAICHE
Neurological Physical and Rehabilitation Medicine Department, Henry-Gabrielle hospital, Hospices Civils de Lyon

Raphael GROSS
MIP UR 4334, University of Nantes, University Hospital of Nantes

Jean-Michel GRACIES
Neurorehabilitation Department, University Hospital Henri Mondor, APHP

Olivier REMY-NERIS
Department of Physical and Rehabilitation Medicine, Brest University Hospital

Nicolas BAYLE
Neurorehabilitation Department, University Hospital Henri Mondor, APHP

Research Article

Keywords: Validity, Reaching, Kinematics, Smoothness, Stroke

Posted Date: January 30th, 2024

DOI: https://doi.org/10.21203/rs.3.rs-3894863/v1
Abstract

**Background:** Movement smoothness is a potential kinematic biomarker of upper extremity (UE) movement quality and recovery after stroke; however, the validity of available smoothness metrics has not been determined in this group. We aimed to measure the responsiveness to change, reliability, convergent and criterion (concurrent and predictive) validity of several smoothness metrics.

**Methods:** This ancillary study of the REM-AVC trial included 31 participants with hemiparesis in the subacute phase of stroke (median time since stroke: 38 days). Assessments performed at inclusion (Day 0, D0) and at the end of a rehabilitation program (Day 30, D30) included the UE Fugl Meyer Assessment (UE-FMA), the Action Research Arm Test (ARAT), and 3D motion analysis of the UE during three reach-to-point movements at a self-selected speed to a target located in front at shoulder height and at 90% of arm length. Four smoothness metrics were computed: a frequency domain smoothness metric, spectral arc length metric (SPARC); and three temporal domain smoothness metrics (TDSM): log dimensionless jerk (LDLJ); number of submovements (nSUB); and normalized average rectified jerk (NARJ).

**Results:** At D30, movements were significantly shorter in duration and trajectory, straighter, faster and smoother (highest effect size for smoothness change: SPARC, 0.76). Intra-subject coefficients of variation were <10% for SPARC and LDLJ, and >30% for nSUB and NARJ. SPARC was strongly correlated with all TDSM, and the TDSM were very strongly correlated with each other. Concurrent validity at D0 was higher for SPARC than TDSM in terms of correlation with proximal UE-FMA, ARAT and index of curvature (r: 0.56, 0.68 and 0.87 respectively, all p<0.01). At D30, concurrent validity was similar between all smoothness metrics. TDSM were very strongly correlated with movement duration at D0 and D30. Finally, SPARC had the highest predictive validity among the four smoothness metrics.

**Conclusions:** Of the four smoothness metrics, the SPARC had the highest sensitivity to change, reliability, construct and criterion validity for the evaluation of the UE in people with moderate to severe stroke in the subacute phase. Among the TDSM, LDLJ was the most reliable.

**Trial Registration:** NCT01383512, https://clinicaltrials.gov/, June 27, 2011

**Highlights**

1. The SPARC is a simple, valid and reliable metric of post-stroke upper limb reaching movement smoothness
2. Validity of the SPARC was higher than temporal domain smoothness metrics, especially in severely impaired individuals
3. The validity of normalized rectified average jerk and number of submovements was unsatisfactory for smoothness assessment in the subacute phase of stroke

Introduction
Spastic paresis of the upper extremity (UE) affects up to 70% of individuals after stroke (1, 2). Three main symptoms are well described in spastic paresis syndrome (3): structural alterations relating to immobility (spastic myopathy, leading to muscle contractures) (4, 5), impaired motor control (stretch-sensitive paresis) of the agonist muscles (6–7), and overactivity of antagonist muscles, (8, 9) (including spasticity (9–11), spastic dystonia (12) and spastic cocontractions (13–16).

Spastic paresis directly alters the movement trajectories and velocity with spatial (poor movement control, less efficient trajectories) and temporal (longer movement duration) discontinuities, resulting in a lack of smoothness (17–19). Changes in the smoothness of the hand trajectory after stroke have been studied during reaching, grasping, and pointing movements (20), and the evaluation of smoothness has been suggested as a valid indicator of the quality of spontaneous motor recovery (21–24) and rehabilitation-induced recovery (19, 25–27).

However, valid and reliable smoothness metrics are needed for the evaluation of changes in movement smoothness in the poststroke spastic paretic UE. To date, many metrics have been used to explore movement recovery after stroke (28). Research involving robotic rehabilitation systems in the last fifteen years has particularly contributed to the development of kinematic metrics, including smoothness, as potential biomarkers for movement recovery (25, 26, 28–30). However, the use of smoothness metrics in clinical research remains limited, as those metrics require particular instrumentation, are often insufficiently defined mathematically (even some robot-specific metrics) and validated, and are often non-reproducible, non-dimensionless (i.e. highly relying on movement time), poorly robust against measurement noise, or are not related to the intermittency of movement (20, 31).

New smoothness metrics that attempt to avoid those limitations have been developed and used to assess point-to-reach and point-to-grasp movement in healthy subjects and individuals after stroke (24, 32–34), namely the log dimensionless jerk (LDLJ), a smoothness metric conceived in the temporal domain and the spectral arc length metric (SPARC). The SPARC was conceived in the frequency domain by Balasubramanian and colleagues who described it as a valid, robust to noise, sensitive, reliable, and practical metric after tests on mathematical models (31, 35).

In an earlier study, we compared the properties of four smoothness metrics currently used in the literature (SPARC, and three temporal domain smoothness metrics (TDSM): LDLJ, number of zero-crossings in the acceleration profile also called number of submovements (nSUB) and normalized average rectified jerk (NARJ)) during UE reaching movements in 32 middle-aged healthy participants (34). The SPARC appeared more reliable, and seemed independent of movement duration in this setting whereas the TDSM were highly time-dependent. A better understanding of the psychometric properties of these metrics is still needed for patients with poststroke UE impairment.

This study aimed to measure the psychometric properties (responsiveness to change, reliability, convergent and criterion (concurrent and predictive) validity) of the SPARC and three TDSM (NARJ, LDLJ and nSUB) for point-to-reach movements in people with moderate to severe impairment in the subacute phase of stroke, before and after a rehabilitation program.
We hypothesized that the three TDSM would be more associated with movement duration (34) while the SPARC would be more associated with movement straightness.

**Methods**

This was an ancillary study of the REM-AVC (Ré-Éducation Mécanisée après Accident Vasculaire Cérébral – Mechanized rehabilitation after cerebrovascular accident) multicenter single-blinded prospective randomized controlled trial, which compared the effects of 20 days (4 weeks, 5 days a week) of self-rehabilitation using a mechanized device with control self-exercises on UE impairment in people in the subacute phase of stroke. More details can be found in the original publication of the study (36). It was conducted in accordance with the Declaration of Helsinki, Good Clinical Practice guidelines and local regulatory requirements (registration number, ID-RCB: NCT01383512, https://clinicaltrials.gov/, registered June 27, 2011), and was approved by the Brest University Hospital Institutional Review Board (n°653). All participants gave written consent to the use of their data.

**Sample**

Of the 218 individuals included in the REM-AVC trial, 37 participants in three centers underwent motion capture of their paretic UE. Six participants were excluded: two did not complete both motion capture assessments and four had uninterpretable data (many artefacts). Among the 31 included, the median (Q1 – Q3) age was 64 (54–72) years and 22 (71%) were males. Twenty-three (74%) participants had experienced an ischemic stroke, 8 a hemorrhagic stroke and the dominant side was affected in 14 (45%) participants. The median (Q1 – Q3) NIHSS score was 11 (7.5–15.5) points and the median (Q1 – Q3) time since stroke was 38 (25–62) days.

**Clinical assessments**

The clinical metrics were the UE-FMA score, the Action Research Arm Test (ARAT), a composite Modified Ashworth Scale (cMAS – the sum of the scores of the elbow flexors and extensors, wrist and finger flexors) and the shoulder passive range of motion (PROM). Each outcome was assessed twice: at inclusion (Day 0, D0) and at the end of the rehabilitation protocol (Day 30, D30). All assessments were performed by a blinded investigator.

**Experimental Set-up**

Participants underwent two 3D motion capture sessions at D0 and D30, during which they performed a reaching task (i.e. reach-to-point) with the impaired UE. Twenty-five reflective markers (14 mm) were placed on UE and trunk anatomical landmarks, by the same investigator at each session, following the International Society of Biomechanics recommendations (37) as illustrated in Fig. 1. Marker trajectories were recorded using a six, eight or nine camera motion capture system (Vicon, MX13 and FX20 camera models, Oxford, UK) at 120 Hz.
Blue: mid-hand marker; red: head of the second metacarpal marker

Participants were seated with their closed fist resting on a table and unconstrained trunk. The shoulder was at 0 degrees of flexion and abduction, and the elbow was flexed at 90 degrees in a neutral pronation-supination position. Participants were asked to reach with their closed fist, at comfortable speed, as close as possible to a single target indicated by a mark on a vertical stick and located in front of them, at 90% of the length of their upper limb and at the clavicle level. The set-up is represented in Fig. 2. The movement was repeated four times, the first attempt being considered as a training movement and thus not recorded. Thus for all participants, a total of 93 movements were recorded and analyzed at each session.

Data analysis

The analyses presented here are focused on the mid-hand marker (placed over the middle of the third metacarpal bone, on the back of the hand). Each recorded trajectory was visually inspected twice by the same investigator to manually define the beginning and end of movements. The beginning of the movement was defined as the first ascending point of the trajectory in an upward direction. The end of the movement was the furthest point of the trajectory in the anteroposterior direction. If large artefacts were observed, the mid-hand marker was replaced by the marker placed over the head of the second metacarpal. Marker position data were computed using WorkStation 5.2.9 (Oxford Metrics, Oxford, UK).

A second order, zero-lag, low-pass Butterworth filter with a 6 Hz cutoff frequency was applied to the trajectories using Python (38) before analyses, except for the SPARC which has an in-built filter (31). The mean value of the three movements was used in the analyses for each outcome. Python was used for all calculations. First, second, and third derivatives of angular displacement (trajectory) of the mid-hand marker data were calculated across all three planes of movement to retrieve the velocity, acceleration, and jerk profiles. Peak velocity and peak acceleration were recorded.

Smoothness was quantified using the SPARC (with $V_{\text{threshold}} = 0.05$ and $\omega_c^{\max} = 20$ Hz as recommended by Balasubramanian et al. (31)) and three temporal domain smoothness metrics (TDSM): NARJ, LDLJ, and nSUB. SPARC and LDLJ are negative metrics (an increase in magnitude towards 0 indicates increased smoothness) whereas NARJ and nSUB are positive metrics (an increase in magnitude indicates a reduction in smoothness). A mathematical description of the metrics is available in a prior publication (34).

The index of curvature (IoC), a measure of movement straightness defined as the ratio of the arc length of the trajectory to the length of the straight line linking the first and the last movement points (39) was calculated. It is reported to approximate movement efficiency in the case of pathological movement (40). The Python code provided by Balasubramanian et al. (31) for computing SPARC and LDLJ was edited to include the calculation of all the kinematic metrics.
Statistics

As participants in both REM-AVC groups received the same amount of treatment and as no demographic, clinical or kinematic variables differed, participants were pooled into a single sample for the purpose of this study. Descriptive statistics were performed to calculate the median values and interquartile intervals. Intrasession variability (between trials) was assessed using the coefficient of variation (CoV\textsubscript{intra}, ratio of the standard deviation to the mean). The normality of data was assessed by visual inspection of the data distribution and a Shapiro–Wilk test. As most data had a non-normal distribution due to visual inspection of the data and sample size, only nonparametric tests were used. Comparisons between D0 and D30 clinical and kinematic variables were performed using Wilcoxon tests, and effect sizes were calculated ([0.1–0.3]: small; [0.3–0.5]: medium; ≥0.5: large).

Responsiveness to change was assessed by comparing (Wilcoxon) smoothness metrics values at D0 then at D30. Median intrasession coefficients of variation at D0 and D30 were used as an estimation of intrasession test-retest reliability. The convergent validity (a subtype of construct validity) of the smoothness metrics was determined by analysing Spearman correlations between the delta values of the kinematic metrics (change observed between D0 and D30), calculated as Δ = D30 value – D0 value. The criterion validity of the smoothness metrics was evaluated on two aspects: 1) concurrent validity using Spearman correlations between smoothness metrics and clinical metrics (UE-FMA and ARAT) and kinematic metrics (movement duration and IoC) at D0 and at D30; and 2) predictive validity using Spearman correlations between smoothness metrics at D0 and clinical and kinematic metrics at D30. Spearman's r was interpreted as weak if < 0.4, moderate if [0.4–0.6], strong if [0.6–0.8] and very strong if ≥ 0.8. All statistical analyses were performed using SPSS v20 (IBM, Armonk, NY).

Results

The only missing demographic, clinical or kinematic data was shoulder range of motion at D0 for one participant. Database for main metrics is available in Appendix A.

Clinical changes

Changes in clinical metrics are presented in Table 1. The UE-FMA total and proximal subscore and ARAT scores were improved at D30 (effect size: large). cMAS and shoulder PROM did not change between D0 and D30.
Table 1
Baseline and final clinical metrics and comparison

<table>
<thead>
<tr>
<th>Clinical measures</th>
<th>Day 0</th>
<th>Day 30</th>
<th>Difference</th>
<th>Effect size</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Median</td>
<td>Q1 – Q3</td>
<td>Median</td>
<td>Q1 – Q3</td>
</tr>
<tr>
<td>UE-FMA</td>
<td>27</td>
<td>19–33</td>
<td>45</td>
<td>33–52</td>
</tr>
<tr>
<td>UE-FMA: proximal subscore</td>
<td>24</td>
<td>17–27</td>
<td>37</td>
<td>26–42</td>
</tr>
<tr>
<td>ARAT</td>
<td>10</td>
<td>3–19</td>
<td>31</td>
<td>19–45</td>
</tr>
<tr>
<td>cMAS</td>
<td>3</td>
<td>1–4</td>
<td>3</td>
<td>1–4</td>
</tr>
<tr>
<td>Shoulder PROM</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anterior flexion</td>
<td>150</td>
<td>110–170</td>
<td>150</td>
<td>120–170</td>
</tr>
<tr>
<td>Abduction</td>
<td>103</td>
<td>90–153</td>
<td>105</td>
<td>90–160</td>
</tr>
<tr>
<td>External rotation</td>
<td>40</td>
<td>15–55</td>
<td>40</td>
<td>20–60</td>
</tr>
</tbody>
</table>

Q: quartile, UE-FMA: upper extremity Fugl Meyer assessment, cMAS: composite modified Ashworth scale, ARAT: Action Research Arm Test, PROM: passive range of motion

a. Change in kinematic metrics, responsiveness and intrasession variability

Changes in kinematic metrics are presented in Table 2. Trajectories and velocity profiles at D0 and D30 are illustrated in Fig. 3. Movements at D30 were significantly shorter in duration and trajectory (less distance covered to reach the target), straighter, faster and smoother according to all four smoothness metrics (medium (mean velocity) to large effect sizes). Peak velocity and peak acceleration did not change significantly. Among the smoothness metrics, SPARC had the smallest intra-individual (between trials) CoV and was the only metric for which CoV was significantly improved at D30 (medium effect size). CoV was also less than 10% for LDLJ but was greater than 30% for NARJ and nSUB.
Table 2
Baseline and final kinematic metrics and comparison

<table>
<thead>
<tr>
<th>Kinematic metrics</th>
<th>D0</th>
<th>D30</th>
<th>Difference</th>
<th>Effect size</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Median</td>
<td>Q1 – Q3</td>
<td>Median</td>
<td>Q1 – Q3</td>
</tr>
<tr>
<td>Duration (s)</td>
<td>2.6</td>
<td>2.1–3.9</td>
<td>2</td>
<td>1.7–2.6</td>
</tr>
<tr>
<td>CoV&lt;sub&gt;intra&lt;/sub&gt;</td>
<td>16</td>
<td>11–25</td>
<td>13</td>
<td>7–24</td>
</tr>
<tr>
<td>Trajectory length (mm)</td>
<td>552</td>
<td>461–709</td>
<td>437</td>
<td>406–623</td>
</tr>
<tr>
<td>CoV&lt;sub&gt;intra&lt;/sub&gt;</td>
<td>7</td>
<td>4–13</td>
<td>4</td>
<td>3–9</td>
</tr>
<tr>
<td>IoC (%)</td>
<td>35.2</td>
<td>22.9–79.2</td>
<td>13.8</td>
<td>10.7–23.1</td>
</tr>
<tr>
<td>CoV&lt;sub&gt;intra&lt;/sub&gt;</td>
<td>25</td>
<td>19–48</td>
<td>35</td>
<td>17–45</td>
</tr>
<tr>
<td>Mean velocity (mm/s)</td>
<td>213</td>
<td>127–273</td>
<td>246</td>
<td>185–319</td>
</tr>
<tr>
<td>CoV&lt;sub&gt;intra&lt;/sub&gt;</td>
<td>13</td>
<td>10–21</td>
<td>13</td>
<td>7–20</td>
</tr>
<tr>
<td>Peak velocity (mm/s)</td>
<td>600</td>
<td>391–721</td>
<td>682</td>
<td>480–745</td>
</tr>
<tr>
<td>CoV&lt;sub&gt;intra&lt;/sub&gt;</td>
<td>12</td>
<td>8–20</td>
<td>11</td>
<td>7–15</td>
</tr>
<tr>
<td>Peak acceleration</td>
<td>2768</td>
<td>1860–3650</td>
<td>3169</td>
<td>1899–4523</td>
</tr>
<tr>
<td>CoV&lt;sub&gt;intra&lt;/sub&gt;</td>
<td>23</td>
<td>11–36</td>
<td>21</td>
<td>13–37</td>
</tr>
<tr>
<td>SPARC</td>
<td>-1.82</td>
<td>-2.14–1.70</td>
<td>-1.61</td>
<td>-1.78–1.51</td>
</tr>
<tr>
<td>CoV&lt;sub&gt;intra&lt;/sub&gt;</td>
<td>8.9</td>
<td>6–14</td>
<td>4.1</td>
<td>2–10</td>
</tr>
<tr>
<td>LDLJ</td>
<td>-10.7</td>
<td>-12.7–9.83</td>
<td>-9.36</td>
<td>-11.0–8.05</td>
</tr>
<tr>
<td>CoV&lt;sub&gt;intra&lt;/sub&gt;</td>
<td>9.1</td>
<td>6–13</td>
<td>7.8</td>
<td>4–13</td>
</tr>
<tr>
<td>nSUB</td>
<td>15</td>
<td>12–28</td>
<td>11</td>
<td>7–17</td>
</tr>
<tr>
<td>CoV&lt;sub&gt;intra&lt;/sub&gt;</td>
<td>33</td>
<td>22–46</td>
<td>32</td>
<td>20–43</td>
</tr>
<tr>
<td>NARJ ×10⁻⁵ (mm/s³)</td>
<td>3.8</td>
<td>1.88–9.71</td>
<td>1.52</td>
<td>0.79–3.31</td>
</tr>
<tr>
<td>CoV&lt;sub&gt;intra&lt;/sub&gt;</td>
<td>48</td>
<td>28–75</td>
<td>38</td>
<td>19–57</td>
</tr>
</tbody>
</table>

Q: quartile, CoV<sub>intra</sub>: median intrasession coefficient of variation, SPARC: spectral arc length metric, LDLJ: log dimensionless jerk, nSUB: number of submovements, NARJ: normalized average rectified jerk

b. Correlations
i. Convergent validity

Correlations between the changes in kinematic metrics between D0 and D30 are presented in Table 3. \( \Delta \text{SPARC} \) and \( \Delta \text{TDSM} \) were moderately to strongly correlated. \( \Delta \text{SPARC} \) was also moderately to strongly correlated with the changes in movement duration and \( \Delta \text{IoC} \). \( \Delta \text{TDSM} \) were very strongly correlated with movement duration, but not with \( \Delta \text{IoC} \). \( \Delta \text{TDSM} \) were strongly to very strongly correlated with each other.

<table>
<thead>
<tr>
<th></th>
<th>( \Delta \text{SPARC} )</th>
<th>( \Delta \text{LDLJ} )</th>
<th>( \Delta \text{NARJ} )</th>
<th>( \Delta \text{nSUB} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \Delta \text{Duration} )</td>
<td>-0.51**</td>
<td>-0.81**</td>
<td>0.90**</td>
<td>0.96**</td>
</tr>
<tr>
<td>( \Delta \text{IoC} )</td>
<td>-0.64**</td>
<td>-0.26</td>
<td>0.35</td>
<td>0.33</td>
</tr>
<tr>
<td>( \Delta \text{nSUB} )</td>
<td>-0.54**</td>
<td>-0.84**</td>
<td>0.89**</td>
<td></td>
</tr>
<tr>
<td>( \Delta \text{NARJ} )</td>
<td>-0.59**</td>
<td>-0.73**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \Delta \text{LDLJ} )</td>
<td>0.67**</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**: p < 0.01, \( \Delta \): D30-D0 value, IoC: index of curvature, SPARC: spectral arc length metric, LDLJ: log dimensionless jerk, nSUB: number of submovements, NARJ: normalized average rectified jerk

ii. Concurrent validity

1. At D0

SPARC was moderately correlated with UE-FMA and its proximal subscore and strongly correlated with ARAT. TDSM were weakly or insignificantly correlated with UE-FMA and moderately correlated with ARAT. SPARC was very strongly correlated with IoC and moderately correlated with movement duration; TDSM were moderately correlated with IoC and very strongly correlated with movement duration. Overall results are presented in Table 4.
Table 4
Spearman correlations (r) between metrics at day 0.

<table>
<thead>
<tr>
<th></th>
<th>UE-FMA</th>
<th>UE-FMAp</th>
<th>ARAT</th>
<th>Duration</th>
<th>IoC</th>
</tr>
</thead>
<tbody>
<tr>
<td>SPARC</td>
<td>0.48**</td>
<td>0.56**</td>
<td>0.68**</td>
<td>-0.59**</td>
<td>-0.87**</td>
</tr>
<tr>
<td>LDLJ</td>
<td>0.39*</td>
<td>0.28</td>
<td>0.46**</td>
<td>-0.89**</td>
<td>-0.57**</td>
</tr>
<tr>
<td>NARJ</td>
<td>-0.36*</td>
<td>-0.23</td>
<td>-0.46**</td>
<td>0.87**</td>
<td>0.56**</td>
</tr>
<tr>
<td>nSUB</td>
<td>-0.30</td>
<td>-0.16</td>
<td>-0.39*</td>
<td>0.93**</td>
<td>0.41*</td>
</tr>
</tbody>
</table>

*: p < 0.05, **: p < 0.01, UE-FMA: upper limb Fugl Meyer score, UE-FMAp: proximal subscore of UE-FMA, ARAT: Action Research Arm Test, IoC: index of curvature, SPARC: spectral arc length metric, LDLJ: log dimensionless jerk, nSUB: number of submovements, NARJ: normalized average rectified jerk

At D30

All smoothness metrics were strongly correlated with UE-FMA and UE-FMAp scores and moderately correlated with ARAT scores. Movement duration was very strongly correlated with all TDSM and only moderately correlated with SPARC. IoC was moderately to strongly correlated with all smoothness metrics.

Table 5
Spearman correlations (r) between metrics at day 30.

<table>
<thead>
<tr>
<th></th>
<th>UE-FMA</th>
<th>UE-FMAp</th>
<th>ARAT</th>
<th>Duration</th>
<th>IoC</th>
</tr>
</thead>
<tbody>
<tr>
<td>SPARC</td>
<td>0.63**</td>
<td>0.66**</td>
<td>0.46**</td>
<td>-0.58**</td>
<td>-0.58**</td>
</tr>
<tr>
<td>LDLJ</td>
<td>0.66**</td>
<td>0.66**</td>
<td>0.47**</td>
<td>-0.82**</td>
<td>-0.62**</td>
</tr>
<tr>
<td>NARJ</td>
<td>-0.74**</td>
<td>-0.73**</td>
<td>-0.55**</td>
<td>0.89**</td>
<td>0.68**</td>
</tr>
<tr>
<td>nSUB</td>
<td>-0.61**</td>
<td>-0.60**</td>
<td>-0.40*</td>
<td>0.81**</td>
<td>0.60**</td>
</tr>
</tbody>
</table>

**: p < 0.01, UE-FMA: upper limb Fugl Meyer score, UE-FMAp: proximal subscore of UE-FMA, ARAT: Action Research Arm Test, IoC: index of curvature, SPARC: spectral arc length metric, LDLJ: log dimensionless jerk, nSUB: number of submovements, NARJ: normalized average rectified jerk

iii. Predictive validity

SPARC at D0 was moderately correlated with UE-FMA, UE-FMAp and ARAT at D30, whereas TDSM at D0 were not correlated with clinical metrics at D30. SPARC at D0 was strongly correlated with IoC at D30, whereas TDSM at D0 were either not correlated (nSUB) or moderately correlated (LDLJ, NARJ) with IoC at...
D30. Weak (NARJ and nSUB) to moderate (SPARC and LDLJ) correlations were observed between smoothness metrics at D0 and movement duration at D30. Time since stroke was moderately correlated at D0 with SPARC (r=-0.44, p < 0.05), but not with TDSM.

Table 5: Spearman correlations (r) between day 0 smoothness metrics and day 30 clinical and kinematic metrics.

<table>
<thead>
<tr>
<th></th>
<th>D0</th>
<th>D30</th>
<th>UE-FMA</th>
<th>UE-FMAp</th>
<th>ARAT</th>
<th>Duration</th>
<th>IoC</th>
</tr>
</thead>
<tbody>
<tr>
<td>SPARC</td>
<td>0.57**</td>
<td>0.54**</td>
<td>0.58**</td>
<td>-0.55**</td>
<td>-0.69**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LDLJ</td>
<td>0.21</td>
<td>0.17</td>
<td>0.31</td>
<td>-0.42*</td>
<td>-0.40*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NARJ</td>
<td>-0.22</td>
<td>-0.18</td>
<td>-0.33</td>
<td>0.39*</td>
<td>0.41*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>nSUB</td>
<td>-0.18</td>
<td>-0.10</td>
<td>-0.26</td>
<td>0.39*</td>
<td>0.26</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

D0: day 0; D30: day 30; *: p < 0.05, **: p < 0.01, UE-FMA: upper limb Fugl Meyer score, UE-FMAp: proximal subscore of UE-FMA, ARAT: Action Research Arm Test, IoC: index of curvature, SPARC: spectral arc length metric, LDLJ: log dimensionless jerk, nSUB: number of submovements, NARJ: normalized average rectified jerk

Supplementary results for correlations between smoothness metrics are presented in Appendix B.

Discussion

This study assessed four smoothness metrics in people with moderate to severe motor impairment before and after one month of intensive physical rehabilitation in the subacute phase of stroke. All smoothness metrics were responsive to change, with the strongest effect size for SPARC. Within session test-retest reliability was poor with nSUB and NARJ. Changes in SPARC were convergent with the changes in the three TDSM (NARJ, nSUB and LDLJ). Before rehabilitation, SPARC was very strongly correlated with movement straightness, and moderately with movement duration whereas the TDSM were very strongly correlated with movement duration, and moderately with movement straightness. Also before rehabilitation, SPARC was more strongly correlated with clinical metrics (UE-FMA and ARAT) than the TDSM. The predictive validity of SPARC was stronger than that of TDSM.

In the present study, most clinical and kinematic metrics improved from D0 to D30. cMAS and shoulder PROM did not change, but these measures are characterized by uncertain validity and sensitivity to change (41–44). Movement duration, trajectory length, straightness and smoothness were abnormal at D0, as is generally observed after stroke (45, 46), and all improved significantly by D30. Kinematic outcomes can provide an accurate indication of UE motor recovery after stroke (47). Participants with high-to-normal UE-FMA still showed deficits in movement kinematic outcomes in a study (39). This was
recently found for the SPARC in a longitudinal study of people with mild stroke (24), and is also suggested by our results in people with more severe stroke. A recent meta-analysis found that smoothness (measured with nSUB) was the most responsive after stroke among few kinematic outcomes (movement duration, peak velocity, shoulder active range of motion (AROM), control strategy, IoC, elbow AROM and trunk AROM) and that it was as responsive to change as the UE-FMA, indicating that clinical and kinematic measures are complementary and provide a comprehensive and accurate follow-up of motor recovery (19). This finding is important for research comparing the effects of different rehabilitation techniques.

We found strong to very strong correlations between the changes in each TDSM from D0 to D30, and a moderate to strong correlation between change in TDSM and change in SPARC. This is in line with data previously reported for theoretical models, healthy individuals and people with Parkinson’s disease (31, 32, 48), and supports the convergent validity of SPARC. The SPARC values during reaching movements of healthy individuals (approximately −1.44 ± 0.02) reported in studies by Engdahl et al. (32), Saes et al (24) and Bayle et al. (34) differ notably from the values found in the present study (D0: -1.82 and D30: -1.61). In addition, the values for the participants with mild stroke in the study by Saes et al (week 1: -1.72, week 5: -1.53) (24) differed from both healthy individuals and the participants with a more severe stroke included in the present study, which supports discriminant validity of SPARC.

At D0, SPARC more strongly correlated with UE-FMA and ARAT scores than did TDSM. However, at D30, the correlations for SPARC and TDSM with UE-FMA and ARAT scores were of similar strength, suggesting that SPARC may have higher concurrent validity than TDSM in the case of more severe impairments. Moreover, SPARC at D0 and change values from D0 to D30 were more strongly correlated with movement straightness than were the TDSM, whereas the TDSM were strongly correlated with movement duration, suggesting stronger construct validity of SPARC than TDSM. Finally, SPARC values at D0 were more strongly correlated with kinematic and clinical measures at D30 than the TDSM, suggesting stronger predictive validity of SPARC. These differences may be explained by the noise-sensitivity and movement duration dependence of TDSM (31), which could have been an issue at D0 as the movements were slower and thus may have generated a higher number of signal artefacts. SPARC was likely also impacted, but to a lesser extent, by the greater heterogeneity of movements at D0 since its CoV_intra was improved at D30. The CoV_intra of NARJ and nSUB were notably higher than those of SPARC and LDLJ, indicating higher test-retest within session reliability for the two latter metrics. Finally, incremental validity (the usefulness of adding movement smoothness to the stroke standard assessment) is yet to be fully determined even if the addition of kinematic movement quantification has been strongly encouraged by an international consensus (49). Overall, the psychometric results of this clinical study complete the mathematical and simulated results of Mohamed Refai et al (20) recommending the SPARC as the only valid smoothness metric in the assessment of reaching tasks after stroke.

Study limitations
Only univariable analyses were conducted owing to the non-normal distribution of the data and the small number of participants. Thus, the correlations, despite their consistency, may be biased by confounding factors. The results may not be generalizable to the people with milder impairments or other types of abnormal movement. The ideal filtering for an optimal noise-to-signal ratio has not been determined for the different smoothness metrics (32). Another choice of filter may have improved TDSM performance, especially at D0. High between-trial variability was observed in the participants with more severe impairment. This particularly impacted nSUB and NARJ which had higher CoVs than LDLJ and SPARC; thus, assessing more UE movements could have led to steadier results. Recording five trials for participants with more severe impairment could be a pragmatic compromise between data robustness and participant fatigue in future studies, as recently suggested (50).

**Conclusion**

SPARC is a valid, reliable and responsive metric for the assessment of smoothness during upper limb reach-to-point movements in the subacute phase of stroke. We recommend using SPARC rather than LDLJ to assess the smoothness of movements of uncontrolled duration. NARJ and nSUB provide less valid and reliable results in this context. The gathering of validity evidence is an ongoing process, therefore future studies using SPARC or other smoothness metrics in different setups should report their findings concerning psychometric aspects.

**Abbreviations**

ARAT: Action Research Arm Test  
cMAS: composite modified Ashworth scale  
CoV: Coefficient of variation  
D0: day 0 (inclusion in the rehabilitation program)  
D30: day 30 (end of the rehabilitation program)  
Δ: D30 value minus D0 value  
IoC: index of curvature  
LDLJ: log dimensionless jerk  
NARJ: normalized average rectified jerk  
nSUB: number of submovements  
PROM: passive range of motion
SPARC: spectral arc length metric
TDSM: Temporal domain smoothness metrics
UE-FMA: upper limb Fugl Meyer Assessment
UE-FMAp: proximal subscore of UE-FMA

Declarations

- Human Ethics and Consent to Participate: This work was conducted in accordance with the Declaration of Helsinki and was approved by the Brest University Hospital Institutional Review Board (n°653). Informed consent was obtained from all participants.
- Consent for publication: Not applicable
- Data availability: The dataset supporting the conclusions of this article is included within the article (and its additional files), more details are available on reasonable request.
- Competing interests: The authors declare that they have no competing interests
- Funding: Funding was provided by the French ministry of health: EMREM_AVC CHU BREST 20 220
- Authors' contributions: BM, ML, ORN, RG, and SB participated in data acquisition; GC and ML analyzed the data; GC, JM, JMG, ML, NB, and ORN interpreted results; GC, JR, LM and NB were major contributors in writing the manuscript. All authors read and approved the final manuscript.
- Acknowledgements: We are very grateful to all participants of this study. We also wish to thank all our colleagues, research therapists and other healthcare professionals at the different study sites for their dedication.

References


47. Thrane G, Sunnerhagen KS, Murphy MA. Upper limb kinematics during the first year after stroke: the stroke arm longitudinal study at the University of Gothenburg (SALGOT). J Neuroengineering Rehabil. 2020 Jun 15;17(1):76.


Figures

Figure 1

Marker placement (left) and virtual reconstruction (right) during motion analysis

*Blue: mid-hand marker; red: head of the second metacarpal marker*

- Midhand marker
- Target
- Trajectory

90% reaching range
Figure 2
Representation of the motion capture set-up representation for the starting position

<table>
<thead>
<tr>
<th></th>
<th>Day 0</th>
<th>Day 30</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vertical Trajectory (mm)</td>
<td><img src="image1" alt="Graph" /></td>
<td><img src="image2" alt="Graph" /></td>
</tr>
<tr>
<td>Vertical Velocity (mm/s)</td>
<td><img src="image3" alt="Graph" /></td>
<td><img src="image4" alt="Graph" /></td>
</tr>
</tbody>
</table>

Figure 3
Trajectories and velocity profiles of day 0 (D0) and day 30 (D30) reaching movements

*Blue line: mean value, lavender: standard deviation area*

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Database.xlsx
- Supplementarymaterials.docx