Concurrent Validity of Stepscan System Measurements

By Patrick Connor, PhD, P.Eng., March 2022 (Last update, May 2022)

Abstract

To evaluate the accuracy of the spatial and temporal measurements and the subsequent calculations of a standard Stepscan system, fast and natural pace walking samples have been collected from 12 individuals. The paper overlay method was used, laying paper out over the Stepscan system such that individuals concurrently leave footprints on paper as they apply pressure to the surface to the Stepscan system. Audio recordings, corrected for the speed of sound, were used to capture footstep timing. In our analyses, we evaluated the following gait metrics: stride length, stride time, cadence, velocity, stride width, and toe out angle. This study confirmed the high degree of accuracy of the hardware and software system achieving ICCs of > .999 for the primary spatial and temporal metrics. The toe out angle and stride width metrics demonstrated small and medium-sized biases, respectively, that are linked to the landmarks and strategy chosen for later computations rather than issues with the hardware or raw recordings.

Introduction

Concurrent validity studies to evaluate parameters of gait on spatial-temporal switch/pressure sensitive mats can be accomplished by collecting data from subjects on 2 or more technologies, one after the other and counter-balanced for order[1-2]. Although this makes it relatively straightforward to perform analyses, and to reference results to existing "gold-standard" systems, each system measures a different set of footprints, albeit from the same individuals, and results are subject to errors in the comparator system as well. A more direct approach records the output of two or more technologies at once so that direct comparisons can be made (e.g., a gait mat and a motion capture system)[3-6]. In the current work, we have chosen this strategy, recording the data from a Stepscan system, while simultaneously marking footprints on paper and recording footstep sounds.

The paper overlay or paper-and-pencil analysis method[3] we used here involves laying paper over a system and collecting the same footprints as seen by the Stepscan system. Audio was selected as a means of capturing the timing of the footsteps because it offers a simple, sub-millisecond sampling approach. To our knowledge, audio footstep sampling has not been used previously in concurrent validity studies of gait mats. Both the paper overlay and audio footstep sampling methods are easily accessible, low cost, and simple in principle – there are no hidden algorithms to contend with, and thus allow clearer comparisons to ground truth. This is in contrast to approaches that compare two complex systems. If the recognized "gold standard" technology used for comparison has the same, possibly erroneous, approach to data recording or calculation of the certain metrics, issues may go unnoticed. Also, it may be more difficult to determine the underlying reasons for any discrepancies between the "test" and "gold standard" systems.

Materials and Methods

Participants

Twelve individuals participated in this study, 6 males and 6 females, ranging from approximately 20 to 50 years old. All participants provided informed consent to take part in our protocol, which was designed in accordance with the standard practices listed in Canada's TCPS2 statement.

Equipment, Setup, and Data Collection Procedure

A Stepscan system is composed of 60 cm x 60 cm pressure sensitive tiles that can be connected in a series to create a pedway or platform of most any practical shape or size. The entire surface area is covered in active sensors at 5 mm intervals, resulting in 14400 sensors per tile, which are scanned at 100 Hz. The sensors are pressure sensitive, which allows the system to capture not only the spatial and temporal parameters of gait, but also underfoot pressures. For the purposes of the present study, however, we will focus on validating only the spatial and temporal parameters of the Stepscan system.

Figure 1 illustrates the setup that was used in this study. An 8-tile, 4.8 m long Stepscan system was setup on a concrete floor. At each end, a tray slightly larger than a tile was placed and a seat behind it. At the center of the system and 50 cm away, a USB microphone (Snowball iCE, Logitech, Model No. A00122) was used to capture audio, sitting 11.5 cm above the surface of the tiles. White paper, 61 cm wide was laid out across the system and clamped in place at each end, the clamps pulling the paper taut. Chalk powder was placed in the trays at either end. Participants were provided two pairs of socks to wear to reduce or prevent transfer of the chalk to their feet. Before beginning a recording session, participants would walk across the system. If their gait was not sufficiently audible, they would be asked to walk more heavily, until their gait could generally be heard well. When ready, participants stood in the chalk tray at one end, covering the base of their feet in the chalk powder. On cue, they would walk across the system, ending in the tray on the opposite end. Here they would wait while the researcher marked the heel (the rear-most edge of the heel) and toe (center of 2nd toe) locations, in case they should be obscured by the second walk. On cue, the participant would then walk back across the system and be seated while the research repeated the marking procedure and prepared a fresh sheet of paper for the next pair of passes across the system. In all, participants each walked 8 times across the system, 4 times at a fast but safe pace followed by 4 passes at a natural pace. During their visit, participants were also given the opportunity to walk across the system unhindered (i.e., without trays, paper, or chalk powder, etc.), have their data processed by the Stepscan system, and a report generated of their gait at a natural pace to take home with them. Data was recorded by the Stepscan clinical software application (Release 2.5.26) and the audio was captured in Audacity (Release 3.1.2), an open-source audio editor. After each pass, the audio was compared to the number of footsteps on the paper and it was trimmed back to include only the footstep sounds corresponding to the steps made on the paper (i.e., not the trays).

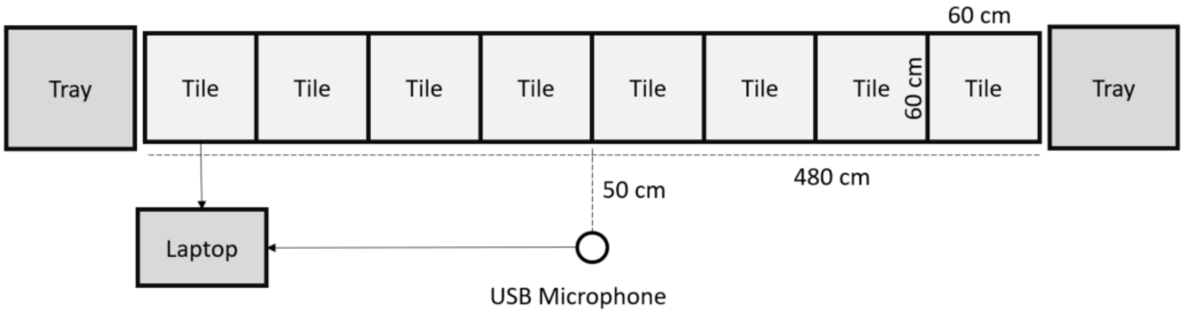


Figure 1, Equipment Setup

Measurements

Because the manual measurements can be very time consuming, the toe and heel locations were marked during the data collection period and the papers were removed and set aside. To facilitate the measurements later on, a similar 8-tile system was setup at table height and the papers were stretched back across these and clamped in place again. The positioning of the paper was similar because of the clamp marks, but not exact to what it was during data collection. A new tape measure (Stanley 5m/16ft) was stretched out across the length of the paper along one side and held in place by the clamp on one end and its metal tab and a piece of tape at the other end. A standard carpenter's square (Mastercraft, 2mm graduations) was guided along the length of the system, allowing us to measure the perpendicular distance to the edge of the system. Thus, the square and the tape measure allowed us to determine the X and Y coordinates of each heel and toe point previously marked. See Figure 2 for a photo of our measurement setup. Between the tools used, the tautness of the paper, and human error each of these measurements, some error is to be expected (see Discussion section below). From the heel coordinates alone, it is possible to calculate stride length. Yet, we chose to also measure the stride length directly using the same tape measure (detached from its placement) to provide a second look. After marking the coordinates on the footprint papers themselves, the numbers were transferred to a smaller diagram derived from the Stepscan software and finally input into a spreadsheet for later comparison.

During data collection, the audio was saved for each participant, as 8 short recordings in series, 1 recording per pass across the system. When the sounds were relatively faint, the audio was sufficiently amplified to help the footstep sounds stand out from the background noise. Determining the initial contact timing was done manually by identifying a region of the signal that was clearly part of the footstep sound and another region of the signal representing the preceding background noise. Then, the transition was located where these two regions met, where there is a shift in the amplitude and or frequency spectrum between them. This transition was used as the initial contact time. In Figure 3, the transition can be clearly seen between the background sound (left) and the footstep (right). A number of footsteps were much less clear, primarily because they were very faintly recorded and so were not as easily distinguishable from the background. Also, some footsteps seemed to have a slight low amplitude frequency change before the primary increase in amplitude occurred. This may have been due to a slight scuff preceding the footstep. The start of this brief putative scuffing sound was usually used as the initial contact time. These were also logged on the associated footprint diagrams and later entered into our data spreadsheet.

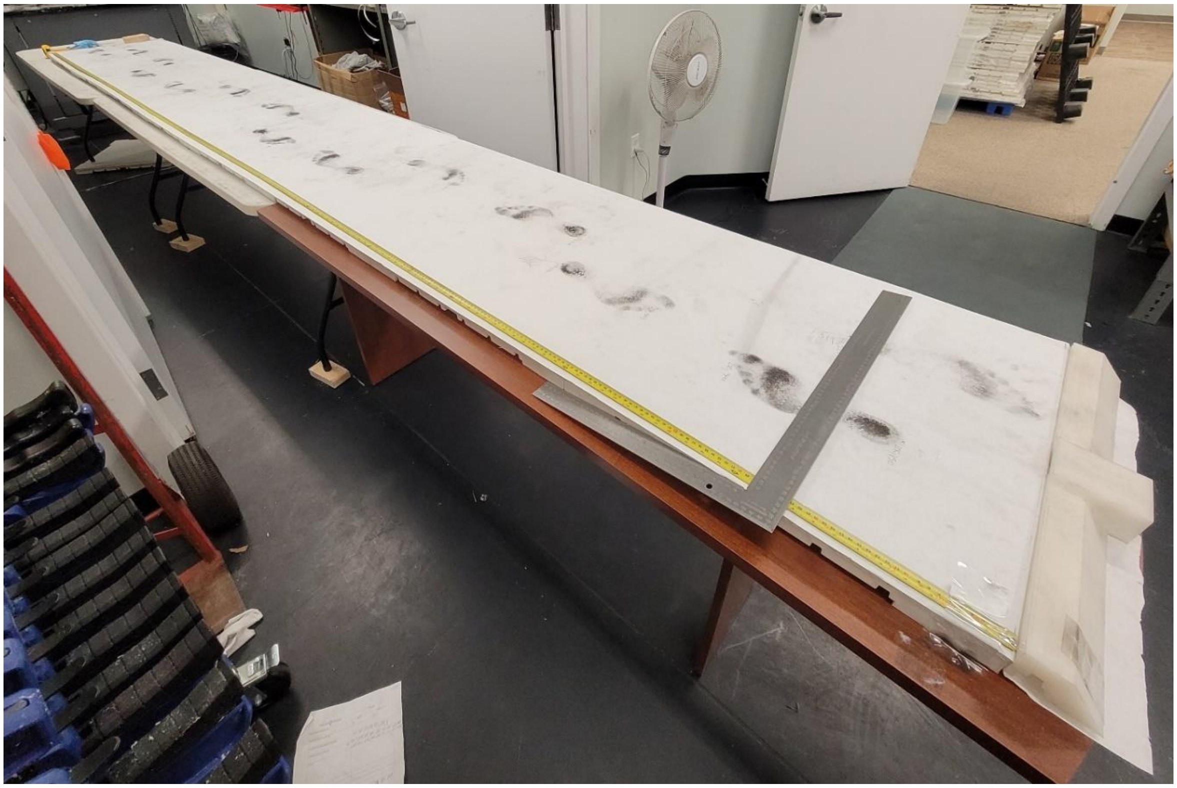
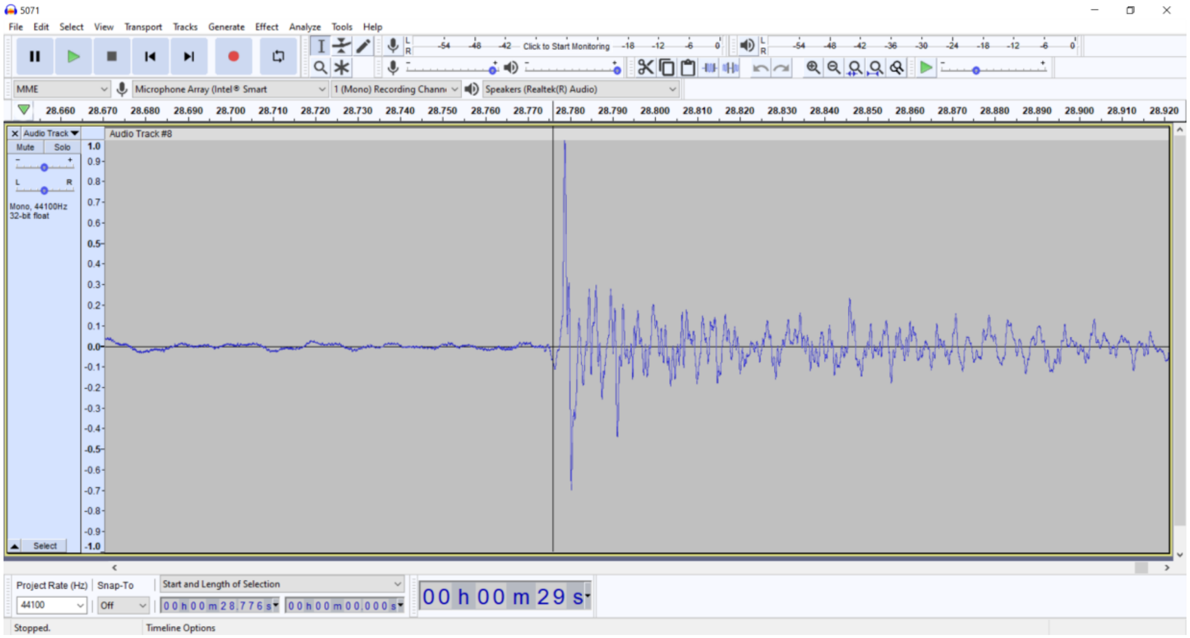


Figure 2, Spatial measurements setup



5071

MME

Microphone Array (Intel® Smart

1 (Mono) Recording Chann

Speakers (Realtek(R) Audio)

Audio Track #8

Mono, 44100Hz
32-bit float

Project Rate (Hz): 44100

Snap-To: Off

Start and Length of Selection: 00 h 00 m 28.776 s 00 h 00 m 00.000 s

00 h 00 m 29 s

Stopped.

Figure 3, Footprint initial contact time measurement by audio

Once the contact times were determined, we used the coordinate location of the associated footprints to subtract the time taken for the sound to travel from the footstep to the microphone via the air. At the more distant locations, the timing due to the speed of sound would otherwise be in error by 7-8 ms.

The data collected by the Stepscan clinical software was processed, a report was generated, and the data of interest was exported to disk in Excel format. Figure 4 depicts the primary processing step in the Stepscan software. Each footprint is automatically classified as left or right, and the direction of walking as left or right. Any changes that need to be made manually are done in this processing step. Screenshots like Figure 4 were used as the diagrams on which we temporarily wrote our physical and audio measures, ensuring a proper matching of the data between Stepscan and the manual measurements.

Analyses

The data from the physical/audio manual measurements and the Stepscan software was collected in parallel in the same spreadsheet, one spreadsheet for the fast pace walking condition and one for the natural pace walking condition.

Each row of data contained heel and toe coordinates of a footprint, as well as the stride length and stride time for the stride ending in said footprint, alongside the stride length and stride times computed by the Stepscan software. The cadence, velocity, stride width, and toe out angle were computed based on our manual measures and were also accompanied by corresponding measures derived from the Stepscan software. In Table 1 below, we briefly explain how each metric was determined from the manual measures and how it was computed by the Stepscan software.

Version: V1
Date: 04-Jan-2022 11:46:24
Note: Source

Figure 4, Stepscan data processing.

Table 1, Compared gait measures

Gait Metric	*Manual Measurement*	*Stepscan Measurement*
Stride Length	Direct measurement between the marked rear heel edges of the first and third feet in a stride.	For the first and third footprints of a stride, their geometric centers are computed. The distance between these centers was then computed as the stride length.
Stride/Gait Cycle Time	Computed as the difference between a stride's first and third footprint initial contact times, as extracted from the audio recordings.	Computed as the difference between a stride's first and third footprint initial contact times.
Cadence	120 / Stride Time	120 / Stride Time
Velocity	Stride length / Stride time	Stride length / Stride time
Stride width	Calculated as the perpendicular distance between the line connecting the first and third *heel coordinates* in a stride and the heel coordinate of the second footstep.	Calculated as the perpendicular distance between the line connecting the first and third footprint *geometric centers* in a stride and the geometric center of the second footstep.
Toe out Angle	Calculated as the angle between the line connecting the heel and 2nd toe coordinates of a footprint (the footprint orientation) and the line connecting the same heel coordinate to the heel coordinate of the footprint two steps ahead (the walking direction).	Calculated as the angle between the mid-line of the footprint (as determined by optimally fitting an ellipse to the shape of the footprint) and the line connecting the geometric center of the same footprint to the geometric center of the footprint two steps ahead (the walking direction).

Some degree of human error was expected in the manual measurements including the transfer of those measurements into our spreadsheets. To catch these, several measures were taken. First, using the heel coordinates for footprints starting and ending a stride, it is possible to calculate stride length. By comparing the directly measured and calculated stride lengths, it was possible to identify gross errors in measurement or transcription. Similarly, computing the X and Y differences between the heel and toe coordinates, additional errors were discovered and resolved. Also, any large differences between Stepscan and manual measurements were investigated, sometimes identifying similar human errors. This quality assurance process does not preclude the existence of error in our manual measures, as will be discussed, but does ensure that no large-scale scribal, typographical, or copying human errors were committed.

Importantly, the Stepscan calculations were performed on the same set of strides and footprints so that differences between the two approaches could be evaluated directly. The Stepscan measures were aligned with our manual measures in the spreadsheets. Differences between them were computed and the average discrepancies (RMSE) were determined. Generally, however, concurrent validity studies tend to compare/contrast the average differences by individual or participant. We report these in terms of intra-class correlation coefficients (ICC(2,k), two-way random, average measures, absolute agreement)[7], which is usually the case for test-retest reliability studies such as this.

Results

Table 2 contains comparisons between the measures described in Table 1. The stride-wise root mean square error (RMSE) measures look at the differences between Stepscan's calculation and the paper overlay / audio measure stride by stride across all participants. In contrast, the participant-wise intraclass correlation coefficients (ICCs) reflect the differences between the two methods, averaging across strides for each participant. Finally, the mean absolute error was computed across participants as well, showing the typical difference between the output of a Stepscan report and the paper overlay / audio method.

Table 2, Stride-wise and Participant-wise Comparisons

Gait Measure	*Stride-wise RMSE (Natural Pace)*	*Stride-wise RMSE (Fast Pace)*	*Participant-wise ICC(2,k)*[*] *(Natural Pace)*	*Participant-wise ICC(2,k)*[*] *(Fast Pace)*	*Participant-wise MAE (Both Paces)*
Stride Length	8.3 mm	7.2 mm	0.99997	0.99993	1.7 mm
Stride/Gait Cycle Time	18 ms	15 ms	0.9998	0.9996	2.5 ms
Cadence	1.6 steps/min	2.0 steps/min	0.9999	0.9997	0.27 steps/min
Velocity	1.6 cm/s	2.2 cm/s	0.99995	0.9991	0.35 cm/s
Stride width	3.9 cm	3.6 cm	0.768	0.607	3.8 cm
Toe out Angle	2.4 degrees	2.4 degrees	0.950	0.911	2.1 degrees

Stride-wise comparisons match direct method-to-method measures. Participant-wise measures result in much less error overall because the plus-and-minus nature of the individual stride-wise errors largely cancel themselves out in the averaging process. The exception is where there are clear positive or negative biases (e.g., stride width and toe out angle), which will be discussed below. Although stride-wise errors are of interest, the participant-wise metrics are the most generally relevant, being the measures that are relied upon in regular research and clinical use.

Discussion

Overall, the results clearly demonstrate that the spatial and temporal measurement capabilities of the Stepscan hardware and software are highly consistent with the paper overlay method, confirming the system's validity. Nevertheless, each measure has some discrepancy shown here and will be discussed. In particular, the stride width and toe out angle values require additional explanation.

To provide context, recall that the Stepscan system has a 5 mm spatial resolution and a 10 ms temporal resolution (100 Hz sampling rate). Thus, the Stepscan system measurements will have a natural lower bound in terms of precision/accuracy. In addition, our paper overlay measurements are not without their errors. First of all, the positions of the heel edge and second toe were not always clearly marked by the chalk powder. Our estimates of these positions could possibly be off by as much as 1 cm in the more extreme cases. In terms of actual human measurement error, one way to evaluate the accuracy of

---

[*] The 95% confidence intervals for stride length, stride time, cadence, and velocity were [0.99, 1.0] at their broadest, while the intervals for stride width and toe out angle spanned [-0.2, 0.95] and [-0.1, 0.99], respectively.

our spatial measurement process (shown in Figure 2) is by comparing the stride length as measured directly from heel to heel coordinates against the calculation of stride length from the associated pair of heel coordinates. The typical difference between these measures was found to be approximately 2.2 mm (RMSE, average of Natural and Fast pace data). Using a Monte Carlo simulation of this scenario, a typical measurement error of +/- 0.85 mm for each of the 5 constituent measurements (two X, two Y, and one point-to-point measurement) would generate this discrepancy. For the temporal measurement process (the audio, see Figure 3), there is no similar way to evaluate the degree of human error. One potential source of error is due to the putative scuff sounds, if they occur in some footsteps and not others. From reviewing the data, they could induce an error anywhere from 5-30 ms on a particular measurement depending on the length of the foot drag/scuff. However, these events were somewhat infrequent, and their total impact is likely low. Instead, it was found that oftentimes the larger temporal discrepancies were found in footsteps at the beginning or end of a pass across the system, where such footsteps were often weakly audible. A weaker audio signal can cause the timing of initial onset of the footstep to be obscured by noise and thus incorrectly determined.

Discrepancies by Metric:

Stride Length – We can estimate the effect of the Stepscan system's spatial resolution by quantizing a pair of supposed footprint locations to the spatial resolution, compute the difference between them and compare this to the actual, unquantized distance between them. Using another simple Monte Carlo simulation under this scenario, the stride length could be expected to be off by 4.17 mm on average on the basis of the spatial resolution alone. Human measurement error as quantified above can account for an additional millimeter. Possible discrepancies in marking the heels of the footprints themselves and using the geometric center of the footprint in the Stepscan calculations (discussed below) likely accounts for the remaining 2-3 millimeters in the average stride-wise differences.

Stride Time – With the Stepscan system's time resolution of 10 ms, a discrepancy of 16.6 ms is expected according to a similar Monte Carlo simulation. This can explain the bulk of the temporal discrepancy. Human error in determining the initial contact times is expected to make up the rest.

Cadence and Velocity – These two metrics are products of the stride length and stride times, so their errors are explained in the same terms as described above.

Stride Width – According to Whittle's "Gait Analysis" text[8], "…walking base (also known as the 'stride width' or 'base of support') is the side-to-side distance between the line of the two feet, usually measured at the midpoint of the back of the heel but sometimes below the center of the ankle joint." Thus this is the definition adopted for the paper overlay method calculation. But there is a discrepancy between this definition and the Stepscan software output that is substantially larger than can be explained by simulation alone. Recall that the Stepscan software uses the geometric center of the footstep, rather than the rear edge of the heel, as the reference points for measuring spatial gait parameters. Although this did not obviously affect the stride length, it did have an impact on stride width, both in the typical stride-wise differences and the participant-wise differences. If the toe out angle were always zero, such that the center of each footprint would be directly above the rear edge point of the heel, the stride widths would be comparable. However, because people typically have a positive toe out, the geometric center of the foot tends to be further lateral than the heel. This can lead to an increase in the stride width of several centimeters more than the usual way of measuring the stride width, which is significant. In fact, all of the participant-wise stride width values from the Stepscan software were larger than those calculated from the paper overlay method. In a way, it would appear that the Stepscan software's approach to computing stride width provides a truer sense of the base of support than the standard stride width measurement. As toe out angle increases, it does widen the usable base of support for increased lateral stability. In fact, a more common definition of "base of support", in contrast to stride width, is the area enveloped by all points of contact with the ground. In research that aims to evaluate correlations between the base of support and some outcome, there might be better sensitivity/specificity to the definition/calculation used by the Stepscan software. Coincidentally, a few studies[9-11] of test-retest reliability involving the GAITRite® system, which calculates stride width using the heel-heel definition, saw the stride width and sometimes toe out angle among their least reliable metrics. Understandably, the standard approach of using the heel points is a more obvious anatomical landmark than the center of the footprint. This, however, is an example of where computerized gait assessment can operate without obvious landmarks and may provide a better approach to measurement. To address these different perspectives, software releases 2.6.9 and beyond will include both the heel-heel and center-of-geometry based stride width calculations, allowing the user to choose their preferred definition. *Update: This change has been implemented and the heel-referenced stride width achieves a participant MAE of 9 mm (down from 38 mm) participant ICCs of .986 and .971 (up from .768 and .607) for the natural and fast speeds, respectively. The associated confidence intervals, however, are relatively broad ([0.2, 1] and [0.02, 1]).*

Toe out angle – The standard approach to measuring toe out angle involves drawing a line from the rear edge of the heel to either the second toe of the foot or simply down its center/midline. The Stepscan approach is slightly different. Stepscan optimally fits an ellipse to the footprint and uses its major (longest) axis as the analogous midline. The typical shape of the human foot, however, leads to the fitting of an ellipse that tends not to have this midline pass through the rear edge of the heel, but rather slightly laterally. The result is that there is a bias of 1-3 degrees less of a positive toe-out angle in Stepscan reports than when calculated from the manual measurements. In fact, the participant-wise Stepscan software toe out angle values were always less than those calculated from our manual measurements. This bias is less substantial than the stride width bias, giving excellent ICCs (> 0.9) overall, albeit with a large 95% confidence interval ([-0.2, 0.95]). To minimize the average effect of this bias, a counter bias of 2 degrees will be included in releases 2.6.9 and beyond of the Stepscan software. *Update: This change has been implemented and aside from subtracting the 2 degrees from the participant-wise MAE, achieved better ICCs of .992 and .986 for the natural and fast speeds, respectively. Even the confidence intervals improved to excellent levels ([0.97,1] and [0.95, 1]) on this same dataset.*

On the whole, the using of a fitted ellipse rather than rear heel edge as a reference point for calculations will have little effect on most gait-related calculations, as evidenced by the low discrepancy in stride length. Aside from stride width and toe out angle, the next most likely affected calculations are expected to be the percentage of medial vs. lateral force applied underfoot, because of the slight difference in which the foot's mid-line is drawn. Yet, the use of ellipse fitting to footprints and geometric centers as a reference point is a natural one and useful for pathological footprints (e.g., toe-walking), and thus may also be used in other software suites.

Finally, in regard to the selected methods for this validation study, the paper overlay method was a good choice. Although the labour involved in extracting the data from the paper was substantial, the human-caused measurement error was quite minimal. This was helped by the nature of the gait measurements being localized to each stride, so we need not be concerned with wrinkles at one end of the paper sheet affecting measurements elsewhere. The use of audio to capture footstep timing was strong in principle. However, having only one microphone meant that footsteps far away were often weak, which could mask the onset of the footstep, and likely introduced error. One possible solution to this problem could be to use a stereo microphone, and spread out the left and right so that they reduce the average distance to the footsteps. The "scuffing" or other sounds captured during recording sometimes also made it difficult to discern initial contact. Yet, even with these challenges, extremely high ICCs (>0.999) were found, even for stride time, which is based solely on comparisons between the Stepscan system and the audio recordings.

Conclusions

This validity study has confirmed the accuracy of the Stepscan hardware and software to capture spatial and temporal parameters of gait with extremely similar participant-wise values to a standard paper overlay method and the audio recordings of footsteps. In the case of calculating stride width and to a lesser degree toe out angle, the Stepscan software employs a slightly different methodology than standard practice, which leads to a noticeable bias in these measures, but few others overall.

References

1. Lynall RC, Zukowski LA, Plummer P, Mihalik JP. Reliability and validity of the Protokinetics Movement Analysis Software in measuring center of pressure during walking. Gait & posture. 2017 Feb 1;52:308-11.
2. Vallabhajosula S, Humphrey SK, Cook AJ, Freund JE. Concurrent validity of the Zeno walkway for measuring spatiotemporal gait parameters in older adults. Journal of geriatric physical therapy. 2019 Jul 1;42(3):E42-50.
3. McDonough AL, Batavia M, Chen FC, Kwon S, Ziai J. The validity and reliability of the GAITRite system's measurements: A preliminary evaluation. Archives of physical medicine and rehabilitation. 2001 Mar 1;82(3):419-25.
4. Bloemendaal MV, Beelen A, Kleissen RF, Geurts AC, Nollet F, Bus SA. Concurrent validity and reliability of a low-cost gait analysis system for assessment of spatiotemporal gait parameters.
5. Walha R, Lebel K, Gaudreault N, Dagenais P, Cereatti A, Della Croce U, Boissy P. The Accuracy and Precision of Gait Spatio-Temporal Parameters Extracted from an Instrumented Sock during Treadmill and Overground Walking in Healthy Subjects and Patients with a Foot Impairment Secondary to Psoriatic Arthritis. Sensors. 2021 Jan;21(18):6179.
6. Vítečková S, Horáková H, Poláková K, Krupička R, Růžička E, Brožová H. Agreement between the GAITRite® System and the Wearable Sensor BTS G-Walk® for measurement of gait parameters in healthy adults and Parkinson's disease patients. PeerJ. 2020 May 22;8:e8835.
7. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. Journal of chiropractic medicine. 2016 Jun 1;15(2):155-63.
8. Whittle MW. Gait analysis: an introduction. Butterworth-Heinemann; 2014 May 12.
9. Van Uden CJ, Besser MP. Test-retest reliability of temporal and spatial gait characteristics measured with an instrumented walkway system (GAITRite®). BMC Musculoskeletal Disorders. 2004 Dec;5(1):1-4.
10. Roche B, Simon AL, Guilmin-Crépon S, Boizeau P, Andriss B, Alberti C, Presedo A, Ilharreborde B, Husson I. Test-retest reliability of an instrumented electronic walkway system (GAITRite) for the measurement of spatio-temporal gait parameters in young patients with Friedreich's ataxia. Gait & Posture. 2018 Oct 1;66:45-50.

11. Menz HB, Latt MD, Tiedemann A, San Kwan MM, Lord SR. Reliability of the GAITRite® walkway system for the quantification of temporo-spatial parameters of gait in young and older people. Gait & posture. 2004 Aug 1;20(1):20-5.