This article may be used for non-commercial purposes in accordance with Wiley Terms and Conditions for Self-Archiving.

The full details of the published version of the article are as follows:

TITLE: Agreement between two inertial sensor gait analysis systems for lameness examinations in horses

AUTHORS: T. Pfau, H. Boultbee, H. Davis, A. Walker, M. Rhodin

JOURNAL TITLE: Equine Veterinary Journal

PUBLISHER: Wiley

PUBLICATION DATE: April 2016

DOI: 10.1111/eve.12400
Agreement between two inertial sensor gait analysis systems for lameness exams in horses

Thilo Pfau¹, Harriet Boultbee¹, Hanna Davis¹, Anna Walker¹, Marie Rhodin²

¹Department of Clinical Science and Services, The Royal Veterinary College, University of London

²Department of Clinical Sciences, Swedish University of Agricultural Sciences, SE-750 07, Uppsala, Sweden

*Correspondence: tpfau@rvc.ac.uk

Keywords: horse, lameness, gait analysis, movement symmetry

Declarations:

Ethical Considerations: The project was granted approval by the Royal Veterinary College’s Ethics and Welfare Committee. Owners of privately owned horses gave signed consent for the use of their horses in the study.

Competing Interests: The authors of this paper have no financial or personal relationships with other people or organisations that could inappropriately influence or bias the content of this paper.

Source of funding: Funding was provided by the Swedish-Norwegian Foundation for Equine Research and the Royal Veterinary College (as part of Harriet Boultbee’s and Hanna Davis’ 3rd year research project).

Acknowledgements: The authors thank Professor Kevin Keegan for constructive criticism on the manuscript and Constanza Gómez Álvarez and Line Greve for help with data collection.
Gait analysis is gaining in popularity for quantification of lameness and two commonly used inertial sensor systems assess trunk movement symmetry: can these be used interchangeably in multi centre studies?

We compared head and pelvic movement symmetry between two inertial sensor gait analysis systems in 13 horses equipped simultaneously with the two systems. The first system quantified dorso-ventral movement in the local reference frame (System A), the other system global vertical movement (System B). Widths of limits of agreement were calculated employing a well-established regression method dealing with systematically changing differences over the range of measured values.

Widths of limits of agreement between system A and system B were narrower for pelvic movement than for head movement. For head movement, they ranged from 6.4 to 6.9mm for in-hand trot and from 7.3 to 9.7mm in the lunge and for pelvic movement from 2.5 to 4.4mm in-hand and from 3.6 to 5.3mm on the lunge.

Widths of limits of agreement between the two investigated inertial sensor gait analysis systems are of comparable magnitude (some equivalent, some marginally higher) to the currently proposed thresholds of 6mm for head and 3mm for pelvic movement used in lameness investigations.

Differences in measurements with two different systems (A and B) obtained from the same horse falling within the reported values should not be seen as a sign of a change in lameness.
Introduction

Equine gait analysis and in particular quantitative assessment of gait parameters in lame horses – e.g. head nod (Buchner et al. 1996) and hip hike (May and Wyn-Jones 1987) – is increasingly performed with systems based on inertial sensors (Keegan et al. 2004, Marshall et al. 2012, McCracken et al. 2012, Starke, et al. 2012, Maliye et al. 2013, Pfau et al. 2014, Rungsri et al. 2014). These systems are based on wireless technology allowing assessment with minimal infrastructure: Inertial sensors mounted on the horse and a laptop computer nearby. Increasing numbers of publications address clinically relevant exercises such as lungeing (Starke et al. 2011, Pfau et al. 2012, Rhodin et al. 2013, Starke et al. 2013, Brocklehurst et al. 2014). In order to avoid misinterpretations of differences between systems in the framework of evidence based medicine, e.g. when a horse is referral to a specialist centre, it is essential to quantify the differences between these two systems. This knowledge is also important for multi-centre research studies when data is collected with different systems. In the context of inertial sensor based systems, potential sources of differences could be related to differences in sensor hardware, the filtering and processing algorithms to derive displacement from the recorded acceleration signals as well as from the selection of strides.

Aim of this study was to quantify the differences between two inertial sensor based gait analysis systems used in clinical practice and research environments under practically relevant conditions. Both systems quantify head and pelvic movement symmetry from inertial sensors mounted over the poll and over the midline of the horse at the level of the tuber sacrale. The first system (System A¹) uses two uni-axial accelerometers mounted on head and pelvis and additionally a uni-axial gyroscope attached to the right forelimb to facilitate identification of stride events. The other system (System B²) uses two inertial sensors each containing one tri-axial accelerometer, one tri-axial gyroscope, and one tri-axial magnetometer per sensor. System A records movement in the direction of the uni-axial accelerometer while system B calculates vertical movement. This difference
is particularly relevant on the lunge when horses lean into the circle (Pfau et al. 2012, Brocklehurst et al. 2014) potentially affecting agreement between the two systems.

Our objectives were (1) to quantify limits of agreement for movement asymmetry quantified with the two systems in trot during in-hand assessment on the straight and during lungeing. We hypothesized, that widths of limits of agreement (Bland and Altman 1986) would be similar to reported values comparing between system B and optical motion capture (Warner et al. 2010) and similar to currently proposed thresholds for system A for the lameness exam: 6 mm for head, 3 mm for pelvic movement asymmetry (McCracken et al. 2012).

Materials and Methods

Animals and facilities

Thirteen horses were recruited from a single riding yard featuring an indoor arena with a sand/fibre based riding surface and an outdoor area with a hard flat surface suitable for trotting horses. The horses were used for a variety of equestrian activities (see table S1) and comprised horses perceived to be sound and well-functioning (i.e. in regular work and in the opinion of their owners not perceived to have performance issues) as well as horses with a previous history of injury (see table S1 for details about horses). The project was approved by the Royal Veterinary College’s Ethics and Welfare Committee.

Data collection

Each horse was simultaneously equipped with two inertial sensor based gait analysis systems. System A\(^1\) comprised of three uni-axial inertial sensors: one uni-axial gyroscope attached to the right forelimb pastern region facilitating stride segmentation and two uni-axial accelerometers mounted over poll and over the midline of the horse at the level of the tuber sacrale to quantify head and
pelvis movement symmetry. Uni-axial acceleration (dorso-ventral) was recorded at 200 Hz with 8 bit digital resolution and over a range of +/-6 times gravitational acceleration (Keegan et al. 2011).

System B² comprised two six degree of freedom-inertial sensors (Pfau et al. 2005, Warner et al. 2010), one mounted over the poll and one over the midline of the horse at the level of the tuber sacrale; both sensors attached immediately behind the corresponding sensor of system A. Sensor data of each of the nine channels (3x acceleration: range +/-18 times gravitational acceleration, 3x angular velocity: range +/-1200 degree/s, 3x magnetic field: range +/-750mGauss) were recorded at 100 Hz and with 16 bit resolution. Data of both systems were transmitted wirelessly from the horse to a nearby laptop computer running the corresponding proprietary data collection software.

Horses were trotted in-hand and while being lunged on a soft equestrian sand/fibre based surface in an indoor riding arena. Lungeing was performed in both directions with a circle radius of approximately 5-7 m. Multiple lungeing trials were acquired for most horses in order to capture stretches of data encompassing steady state locomotion (horse moving at consistent speed and circle radius; judged subjectively). Data collection was manually started and stopped at approximately the same time for the two inertial sensor systems. Data collection was continued until at least 25 strides of steady state locomotion – judged subjectively by the experimenters at the time of data collection – were available for each of the exercise conditions.

Data processing

Data were processed with the corresponding software packages for each system. For both systems this procedure comprised filtering, stride segmentation and double integration from acceleration to displacement (Keegan et al. 2001, Keegan et al. 2004, Pfau et al. 2005, Warner et al. 2010, Starke, et al. 2012b). Sensor based (System A) or vertical (System B) displacement values calculated over each stride cycle were then used to determine movement symmetry for each stride cycle. Movement symmetry was characterized in both systems by calculating the differences in minimum and
maximum head and pelvic displacement that occurs during and after stance of right and left halves of each full stride cycle (HDmin, HDmax, PDmin, PDmax, (Kramer et al. 2004, Keegan et al. 2011), or MinDiff, MaxDiff, (Starke et al. 2011)). For each horse a mean value for all strides was calculated for each exercise condition (i.e. for each trial representing one of the exercise conditions, trot in straight line or on left or right rein). Prior to further statistical analysis values for HDmax and PDmin for system A were multiplied by -1 to match the sign convention of system B: positive values for MinDiff for left hind and right forelimb lameness and for MaxDiff for right hind and left forelimb lameness, negative values for the MinDiff for right hind and left forelimb lameness and for MaxDiff for left hind and right forelimb lameness.

**Data analysis**

Head and pelvic movement symmetry measures were compared between the systems based on procedures described previously for method comparison studies (Bland and Altman 1986). Averages of and differences between the mean symmetry values quantified for the two systems were calculated for each trial. Evaluation of scatter plots of the difference values over the mean values (Bland and Altman, 1986) showed that differences between the two systems were systematically affected by the measurement value: increasingly negative differences with increasing mean value. As a consequence a published regression method (Bland and Altman 1999) was employed to take into account this systematic difference when establishing widths of limits of agreement values for each symmetry measure. In brief, rather than calculating limits of agreement that are constant over the range of measurements (Bland and Altman, 1986), non-constant estimates of mean difference and upper and lower limits of agreement are calculated based on regression (Bland and Altman, 1999). In order to facilitate calculation of matched movement symmetry values for the two systems, taking into account the identified systematic differences between the two systems, we also fitted linear regression lines to scatter plots of system A values versus system B values and present slope and intercept of these.
Results

A total of 81 trials were successfully recorded for 12 out of the 13 horses providing mean movement symmetry values for >25 strides per trial for both systems. Operator error during data collection prevented use of the data of one horse for further analysis.

Limits of agreement

Figure 1 illustrates the limits of agreement established by the regression method (Bland and Altman 1999) showing both the mean difference between the two systems and the upper and lower limit of agreement (mean difference +/- 2 SD of differences) over the range of observed movement symmetry measures for each parameter. All four show a systematic difference between the systems, indicating a decrease in difference with increasing symmetry value. Width of limits of agreement values (difference between upper and lower limit) are smaller for the pelvic measures (3-5mm, Table 1, ‘all’) than for the head symmetry measures (7-9 mm, Table 1, ‘all’).

Differences between straight line and lungeing

In order to establish whether agreement was different for straight line and lungeing, we analysed the widths of limits of agreement values separately for the straight-line and for the lungeing trials. This is illustrated in Figure 1 with different colors (blue: lunge; red: in-hand). As a consequence of the systematic differences the widths of the limits of agreement varied as a function of the measured value, i.e. the lines illustrating upper and lower limit in figure 1 are not parallel to the x-axis. In order to present one representative value for the agreement per condition an average value for the widths of limits of agreement was calculated over a range of symmetry values from -20 mm to +20mm; this range covers more than 95% of the movement symmetry values measured in this study. Table 1 shows that the widths of limits of agreement vary between +/- 2.5 mm (PDmax, straight line) and +/- 9.7mm (HDmax, lungeing) with smaller values for pelvic movement (+/- 2.5mm to +/- 5.3mm) than for head movement (+/- 6.4mm to +/- 9.7mm). Average limits of +/- 6.2mm were quantified.
across all four measures for the combined data set, of +/- 6.5mm for lungeing and +/- 5.1mm for straight line trials. Finally, table 2 presents slope and intercept values for linear regression lines fitted to scatter plots of system A versus system B values. The presented values allow calculation of system B values from system A values.

Discussion

Here we have compared two commonly used inertial sensor based equine gait analysis systems that quantify head and pelvic movement symmetry and we established limits of agreement after correcting for systematic differences between the systems (Bland and Altman 1986, Bland and Altman 1999).

A ‘worst case scenario’ study design was chosen to reflect the practical scenario we have in mind: a horse gets transferred between veterinarians, e.g. from first opinion practice using one system to a specialist referral centre using the other system. The question is then when comparing movement symmetry values whether the horse shows an improvement, a worsening or no change.

Synchronization between the systems was hence implemented by recording approximately the same series of strides with each system by starting and stopping the recording simultaneously (no hardware synchronization) since in the above scenario, no information about the selected strides will be available. With this ‘worst case scenario’ approach we have—in our opinion—achieved promising results.

Studying the scatter plots in Figure 1 (Bland and Altman 1986) and slope values presented in table 2 between system A and system B values it becomes apparent that system A consistently underestimates the amount of movement asymmetry compared to system B: decreasing differences with increasing asymmetry values. System B has been shown previously to marginally overestimate displacement compared to an optical motion capture system by 0.7 to 2% ([Warner et al. 2010], table 1). Here, the slopes of the regression lines of figure 1 as well as the slope values found in table
suggest that system B overestimates asymmetry by considerably more than 2% compared to system A. As a consequence, system A would underestimate the ‘true’ amount of asymmetry were it to be compared to motion capture.

By using the suggested regression method (Bland and Altman 1999) it was possible to take into account the systematic differences between the two systems and to establish average width of limits of agreement values across a range of movement symmetry values. We chose to calculate the widths of limits of agreement over a range of symmetry value from -20mm to +20mm hence including over 95% of the values presented here during trot in-hand and on the lunge for a range of horses with/without history of musculoskeletal problems (table S1).

By comparing the widths of limits of agreement (the 95% confidence interval) to threshold values used in the context of the clinical lameness exam (McCracken et al. 2012), it is possible to make a judgement about the interchangeability of measurements for the task of classifying horses into ‘sound’ and ‘lame’. If the disagreement is higher than the thresholds, then a classification into sound and lame will statistically result in discrepancies in more than 5% of cases. In a study with system A, thresholds of 6 mm for HDmin or HDmax and 3 mm for PDmin or PDmax have been presented (McCracken et al. 2012). The widths of limits of agreement values of +/-8.8 mm and +/-7.2 mm for HDmax and HDmin and of +/-3.4 mm and +/-5.2 mm for PDmax and PDmin observed across all exercise conditions (Table 1) suggest that the limit of agreement values are equivalent (PDmax) or marginally outside (all others) these threshold values. However the widths of limits of agreement values on the straight (Table 1) are with +/-6.9 mm and +/-6.4 mm for HDmax and HDmin and +/-2.5 mm and +/-4.4 mm for PDmax and PDmin closer (HDmax, HDmin, PDmin) or even marginally below (PDmax) these thresholds and it should be emphasized that the thresholds (McCracken et al. 2012) have been defined based on straight line trot.

The widths of limits of agreement values are also similar (slightly larger for head movement and slightly smaller for pelvic movement) to the values presented previously for a comparison between
system B and an optical motion capture system (+/-4 and +/-8mm, (Warner et al. 2010)). The two inertial sensor systems hence agree similarly well than system B with the optical system. This seems interesting to note since – in contrast to the earlier study (Warner et al. 2010) where exact synchronization between inertial sensors and motion capture was performed – here with our practical ‘worst case scenario’ approach we made use of the automated stride selection provided by the different inertial sensor software packages. The influence of exact time synchronization when comparing between different inertial sensor systems should be further investigated.

An additional source of ‘mismatch’ between the two systems is the physical location; here we placed system B sensors directly behind the corresponding system A sensor (approximately 0.05m between sensors). Only limited variation has been documented from inertial sensor measurements placed along the spine (Warner et al, 2010) with inter-sensor distances of approximately 0.15 to 0.2m. Abaxial misplacement of motion capture markers (Starke et al, 2012c) has been shown to have more influence on movement symmetry measurements (up to 11mm when misplaced by 0.07mm; typical inter- and intra-operator variation in marker placement has been reported to be considerably less than 0.07m (Weller et al, 2006)). Care should hence be taken to place sensors in the midline of the horse since the sensing elements (in particular relevant here accelerometers, gyroscopes) may not be in the centre of the physical sensor housing.

The horses used in this study varied in breed, age, sex, use and presence/history of musculoskeletal problems (see table S1). In the context of the study design employed here – comparing two gait analysis systems simultaneously mounted on the same horse – the variability between horses is not a disadvantage since comparisons are made within horses. On the contrary, if all horses had been completely sound, i.e. showing symmetrical movement on the straight and only small asymmetries on the lunge (see e.g. (Starke et al. 2011, Pfau et al. 2012)), then the comparison between the systems would likely have covered a much smaller range of values (x-axis in Figure 1). Estimates of the limits of agreement would then only have been applicable for the small range of values observed
in sound horses. Here average values of movement symmetry measures between the two systems cover a range of approximately +/-25 mm for head movement and of up to +/-15 mm for pelvic movement (x-axis, Figure 1). This is similar to what has been reported previously for similar lungeing conditions (Pfau et al. 2012) and we have used a similar range of +/-20 mm to calculate widths of limits of agreement from the regression approach, which covers more than 95% of the symmetry values in this study.

Conclusions

After regression based correction for systematic differences between the two systems, the widths of limits of agreement values for comparison of straight line trials are within or marginally outside currently proposed thresholds of detecting lameness in horses (6 mm for head movement, 3 mm for pelvic movement). Differences in measurements between the two systems obtained from the same horse that fall within the widths of limits of agreement values reported here should not be seen as a sign of a change in movement symmetry in this horse. These are +/-6.9 mm and +/-6.4 mm for HDmax and HDmin and +/-2.5 mm and +/-4.4 mm for PDmax and PDmin on the straight. On the lunge these are +/-9.7 mm and +/-7.3 mm for HDmax and HDmin or +/-3.6 mm and +/-5.3 mm for PDmax and PDmin.

Manufacturers’ addresses

1 LamenessLocator, Equinosis, LLC, Columbia, Missouri, United States of America

2 MTx, Xbus system, Xsens Technologies B.V., Enschede, The Netherlands

References


**Supplementary Information Items**

**Table S1:** Information about horses participating in study.
Figure 1: Difference between system A and system B symmetry measures (A-B, y-axis) as a function of average value of both systems ((A+B)/2, x-axis) for each of the 81 trials (red: straight line trials; blue: lungeing trials) from the 12 horses for which data was successfully recorded. The widths of the limits of agreement are illustrated by the green lines including +/- 2 SD of difference values over the range of observed movement symmetry values.

A: difference in head movement minima (HDmin), B: difference in head movement maxima (HDmax), C: difference in pelvic movement minima (PDmin), D: difference in pelvic movement maxima (PDmax).
Tables:

**Table 1**: Width of limits of agreement (+/-2*SD of differences) established as an average across the range of -20 mm to +20 mm for average values between system A and B measurements (x-axis in Figure 1) for all four symmetry measures for data from all trials, for straight-line trials and for lungeing trials. For all four symmetry measures the widths of the limits of agreement are narrower for straight-line trot than for lungeing. The difference between the widths of the limits of agreement of straight line and lungeing is also given (lunge-straight). All values are given in mm.

<table>
<thead>
<tr>
<th></th>
<th>All</th>
<th>lunge</th>
<th>straight</th>
<th>lunge-straight</th>
</tr>
</thead>
<tbody>
<tr>
<td>HDmax</td>
<td>+/- 8.8</td>
<td>+/- 9.7</td>
<td>+/- 6.9</td>
<td>2.8</td>
</tr>
<tr>
<td>HDmin</td>
<td>+/- 7.2</td>
<td>+/- 7.3</td>
<td>+/- 6.4</td>
<td>0.9</td>
</tr>
<tr>
<td>PDmax</td>
<td>+/- 3.4</td>
<td>+/- 3.6</td>
<td>+/- 2.5</td>
<td>1.1</td>
</tr>
<tr>
<td>PDmin</td>
<td>+/- 5.2</td>
<td>+/- 5.3</td>
<td>+/- 4.4</td>
<td>0.9</td>
</tr>
<tr>
<td>average</td>
<td>+/- 6.2</td>
<td>+/- 6.5</td>
<td>+/- 5.1</td>
<td>1.5</td>
</tr>
</tbody>
</table>

**Table 2**: Slope (a) and intercept (b) values of the regression equations for calculation of system B movement symmetry values (y) based on system A movement symmetry values (x) for straight line ('in-hand') and lunge trials.

Equation used: \( y = a \times x + b \) based on sign convention for system B (see materials and methods for details).

<table>
<thead>
<tr>
<th>symmetry measure</th>
<th>straight</th>
<th>lunge</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>slope a</td>
<td>intercept b</td>
</tr>
<tr>
<td>HDmin</td>
<td>1.2204</td>
<td>0.4118</td>
</tr>
<tr>
<td>HDmax</td>
<td>1.2126</td>
<td>2.3728</td>
</tr>
<tr>
<td>PDmin</td>
<td>1.2218</td>
<td>1.7242</td>
</tr>
<tr>
<td>PDmax</td>
<td>1.4545</td>
<td>1.8118</td>
</tr>
</tbody>
</table>