Response to Reviewer 1 Comments

Point 1: The manuscript presents a feasibility study of a sensor platform in assessing gait and sleep. It is well written and the results are promising. There are a few issues to be addressed before the paper can be accepted.

Response 1: We thank the reviewer for his/her positive comment and providing us with feedback.

Point 2: It is mentioned in the abstract that "... Clinical usage and implementation are of utmost relevance and have not been studied." Actually, D. Johansson et al., published a review article to summarize all clinical applications of wearable sensors in epilepsy, PD, and stroke in 2018. https://link.springer.com/article/10.1007/s00415-018-8786-y The authors need to do more literature work before making this statement. Also, if the authors can not defend this, the innovation of this work will be questioned.

Response 2: We thank the reviewer for pointing out the review of Johansson et al. (Johansson et al.). This systematic review aggregated both quantitative and qualitative knowledge from clinical research with wearable sensors. Not entirely consistent with the clinical research studies in this review, we focused on continuously monitoring outside the research setting by using a self-regulated system in clinical practice. Moreover, in the review of Johansson et al., it was addressed that there are some challenges that need to be met before wearables can be integrated in the clinical practice including acceptable integration in daily life, confidence in technology, and the need to consider individualization. This is indeed what our study addressed; we tried to investigate the integration in 'daily life on the clinical ward' and to show that technology can be used in the clinical setting. These studies will help to increase the acceptability and confidence in technology by the medical personnel.

With respect to the review of Johansson et al., half of the 12 mentioned studies reported on missing and incomplete data were performed in clinical setting. Furthermore, these clinical studies measured for a short period (8 – 66.8 hours) and/or without a self-regulated system (the sensors tested had to be renewed every 2 days or data export was not automated). In contrast, our study focused on continuously monitoring for a one week period and let the clinical setting run its course on the rehabilitation ward for patients with an incomplete spinal cord injury or stroke. Providing a more realistic representation of the actual inpatient conditions in which the system eventually could be used, monitoring patients over a longer period of time.

We also reviewed the scientific literature regarding implementation of other wearable sensors (vital sensors, glucose-, behavioural monitors) in clinical practice. Still many wearable-based studies are conducted in controlled environments and/or are of short duration (Weeks et al.; Kroll et al.; Levitt et al.; Cameron et al.; Kleiman et al.). Those studies have reported on missing data in the same range of the studies already mentioned in our discussion section.

Hence, studies regarding implementation of sensor technology on the ward with an continuous, automated self-regulating system are important and to our knowledge scarce. With the words "clinical usage" in the abstract we were referring to clinical practice outside a research setting.

Actions 2: For clarification we have rewritten this sentence and it reads now (L18): *"Implementation in clinical practice is of utmost relevance and has scarcely been studied."*

In the introduction we have added the review of Johansson as a reference to the factors involved for implementation of sensor technology in clinical practice (L77).

In the discussion, we changed the loss of data percentage range from 21% to 24% based on the review of Johansson et al (Johansson et al.) (L284). We have added implementation studies of other sensors in other fields and rephrased two sentences to: *"Studies focusing or reporting on missing data are scarce. However, within controlled standardized settings or inpatient studies of short duration a minimal loss of data has been reported ranging from 0 up to 24% when using various sensor technologies (Mansfield et al.; Weeks et al.; Weenk et al.; Johansson et al.; Kleiman et al.; Kroll et al.; Levitt et al.; Cameron et al.)."*

Based on the comments of reviewer 2, we also added the following in the discussion regarding the technical and human interaction error causes (L304-L315): *"Missing data resulting from technical errors was 10%, which matches with previous literature (Johansson et al.). Frequently described technical errors in various wearable-sensor studies reporting on missing data are: battery failure, device failure, storage problems or connectivity issues (Beniczky et al.; Bayly et al.; Hale et al.; Weiss et al.; Askim et al.; Uswatte et al.; Velez et al.; Prajapati et al.; Weenk et al.; Sanchez et al.). The technical errors encountered in our study are in accordance with these findings. However, the decrease of technical errors over time indicates that technical hard- and software errors could be solved and the platform was optimized during the iterative process. In contrast, the number of human interaction errors did not significantly change over time. Missing data from human interaction errors was 2% which is less compared to previous literature (Johansson et al.). A system with human and technology interaction is prone to errors. Not wearing, losing, and incorrect use of sensors are often reported human interaction errors (Bayly et al.; Askim et al.; Weiss et al.; Gebruers et al.; Uswatte et al.; Urbin et al.; Mudge and Stott; Hale et al.; Velez et al.; Weenk et al.) and were also reported in our study. However, those errors can be reduced to a minimum with proper management."*

Additionally two sentences are combined into one sentence as follows (L321-L324): *"If this is the case this is more in line with literature, which reported a majority of missing data related to human interaction errors in contrast to a minority related to technical difficulties (Mansfield et al.; Johansson et al.)."*

Point 3: In section 3.3 Error causes, "Over time, the number of technical errors decreased, whereas no trend was seen in the number of human interaction", this claim is not supported by any data, figure, or table.

Response 3: We agree with the reviewer that there is no data supporting our claim. We based our claim on the trend as shown in Figure 3. To support our claim, we have adjusted the stacked bars in Figure 3 to single bars for each type of error (human, technical or unknown) to have a better presentation of development of technical and human interaction errors over time. Moreover, we performed a linear regression analysis between number of errors (human and technical) as dependent variable and the number of participants as an independent variable. Regression analysis revealed for the technical errors a negative significant correlation (R^2=.24, p=.016) with a slope of -0.16 and intercept of 5.8. For the human error, there was no significant correlation (R^2=.11, p=.11) with a positive slope of 0.07.

Actions 3: The following sentence was added to the methods: *"To determine the development of errors over time, a linear regression analysis was performed between the number of errors (technical and human interaction) as dependent variable and number of participants as independent variable." (L181-L183).*

In the result section we have rephrased and added the following (L224-L228): *"Over time, the number of technical errors significantly decreased, whereas no significant increase was seen in the number of human interaction errors. A significant negative correlation (R^2=.24, p=.016) with a slope of -0.16 and*

intercept of 5.8 was found for the technical errors. For the human interaction errors, there was no significant correlation (R^2=.11, p=.11) with a positive slope of 0.07."

In the discussion we have rephrased two sentences to (L307-L310): *"However, the decrease of technical errors over time indicates that technical hard- and software errors could be solved and the platform was optimized during the iterative process. In contrast, the number of human interaction errors did not significantly change over time."*

Figure 3 and its legends has changed to (L230-239):

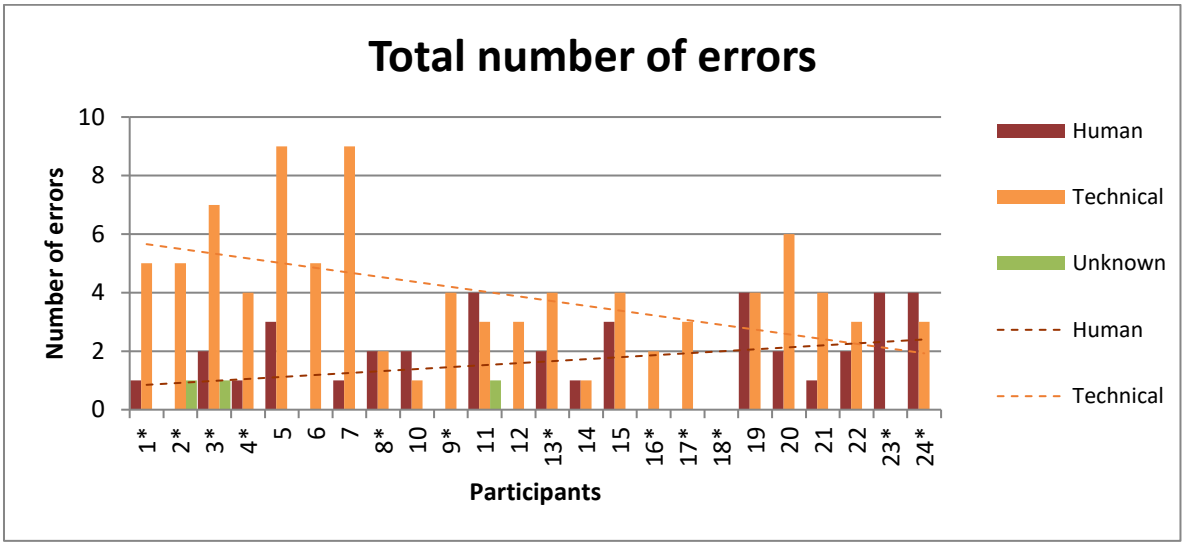

Figure 3. Number of technical, human interaction, and unknown errors for each participant. Participants are ordered chronologically based on inclusion date. Incomplete Spinal Cord Injury (iSCI) patients are marked with an asterisk (*), others are stroke patients. Linear regression lines based on errors over time have a significant negative correlation (R^2=0.24) for technical errors and non-significant positive correlation (R^2=0.11) for human interaction errors.

Point 4: Regarding the missing data discussion, the manuscript presents a few studies to compare the percentage of missing data and claims it will be sufficient if more than 2 days of data is acquired. However, every study is different from others in terms of study design and study goals. 2 days of data cannot be used as a universal criterion. The point is to validate if the missing data has a key information component or not. Thus, this discussion cannot convince the readers, at least the reviewer, the missing data is not important.

Response 4: We agree with the reviewer that 2 days of data cannot be used as a universal criterion and that it depends on which parameters you want to address. Indeed, we have to validate if the missing data has a key information component or not. We based our claim on the literature. Ishikawa et al. (Ishikawa et al.) found a small between-day variability over the week, leading to the interpretation that within a week any 2-day combination of step count values can be averaged to obtain a stable value of step activity in adults with iSCI. For step count, another study (Mudge and Stott) suggest that a three day monitoring period appears to be at least as reliable as seven-day monitoring. Because our main interest is in gait, we concluded that a whole complete measurement week is not necessarily needed and less days might already be sufficient. However, when focusing on physical activity, data collection over 7 days has been proposed (Hale et al.; Matthews et al.). Therefore, we agree with the reviewer that it depends on the outcome parameter if the missing data has a key component or not.

Action 4: Since we did not look at the content of the data, we have refined our statement in the discussion and it now reads (L291-L299): *"The amount of data required to measure these changes depends on the outcome parameter. A minimum of two to three days within a week has been shown to be necessary to reliably measure step count (Ishikawa et al.; Mudge and Stott). However, for physical activity, data collection over 7 days is proposed (Hale et al.; Matthews et al.). Depending on the information of interest, a complete measurement week is not necessarily required and less days might already be sufficient. In the current study, all participants except one had more than two days of obtained data. Moreover, 13 of the 24 participants had no missing days at all and 5 had only 1 missing day over a week which is sufficient for certain gait outcome parameters."*

Point 5: Patients questionnaire terms are not well explained, e.g. "cognitive load" is not a self-explainable term.

Response 5: The term "cognitive load" was used as an overarching term comprising the variety of psychological-related given answers by the participants. The term "cognitive load" was explained in the result section with examples (L258-L260): the fear of forgetting to put on the IMUs and for paying attention to correct placement of the IMUs, the idea that something is around the ankles, and asking nursing staff for help. However, the term "cognitive load" was already mentioned earlier in the result section.

Action 5: We added the following two examples in L198: *"e.g. too much to think about and the idea that they have to wear something around the ankles".*

Point 6: There are also some typos and missing spaces in the paper

Response 6: The manuscript has been carefully checked and corrected.

References

Askim, Torunn, et al. "Changes in Physical Activity and Related Functional and Disability Levels in the First Six Months after Stroke: A Longitudinal Follow-up Study." *Journal of Rehabilitation Medicine*, vol. 45, no. 5, 2013, pp. 423–28, doi:10.2340/16501977-1137.

Bayly, Jade, et al. "Time-Frequency Mapping of the Rhythmic Limb Movements Distinguishes Convulsive Epileptic from Psychogenic Nonepileptic Seizures." *Epilepsia*, vol. 54, no. 8, Aug. 2013, pp. 1402–08, doi:10.1111/epi.12207.

Beniczky, Sándor, et al. "Detection of Generalized Tonic-Clonic Seizures by a Wireless Wrist Accelerometer: A Prospective, Multicenter Study." *Epilepsia*, vol. 54, no. 4, Apr. 2013, pp. e58–61, doi:10.1111/epi.12120.

Cameron, Faye M., et al. "Closed-Loop Control Without Meal Announcement in Type 1 Diabetes." *Diabetes Technology & Therapeutics*, vol. 19, no. 9, 2017, pp. 527–32, doi:10.1089/dia.2017.0078.

Gebruers, N., et al. "Actigraphic Measurement of Motor Deficits in Acute Ischemic Stroke." *Cerebrovascular Diseases*, vol. 26, no. 5, Nov. 2008, pp. 533–40, doi:10.1159/000160210.

Hale, Leigh A., et al. "Measuring Free-Living Physical Activity in Adults With and Without Neurologic Dysfunction With a Triaxial Accelerometer." *Archives of Physical Medicine and Rehabilitation*, vol. 89, no. 9, Sept. 2008, pp. 1765–71, doi:10.1016/j.apmr.2008.02.027.

Ishikawa, Saori, et al. "Reliability of Daily Step Activity Monitoring in Adults with Incomplete Spinal Cord Injury." *Journal of Rehabilitation Research and Development*, vol. 48, no. 10, 2011, pp. 1187–94, doi:10.1682/JRRD.2010.09.0190.

Johansson, Dongni, et al. “Wearable Sensors for Clinical Applications in Epilepsy, Parkinson’s Disease, and Stroke: A Mixed-Methods Systematic Review.” *Journal of Neurology*, vol. 265, no. 8, Dr. Dietrich Steinkopff Verlag GmbH and Co. KG, Aug. 2018, pp. 1740–52, doi:10.1007/s00415-018-8786-y.

Kleiman, Evan, et al. “Using Wearable Physiological Monitors with Suicidal Adolescent Inpatients: Feasibility and Acceptability Study.” *Journal of Medical Internet Research*, vol. 21, no. 9, 2019, doi:10.2196/13725.

Kroll, Ryan R., et al. “Use of Wearable Devices for Post-Discharge Monitoring of ICU Patients: A Feasibility Study.” *Journal of Intensive Care*, vol. 5, no. 1, Journal of Intensive Care, 2017, pp. 1–8, doi:10.1186/s40560-017-0261-9.

Levitt, David L., et al. “Insulin Pump and Continuous Glucose Monitor Initiation in Hospitalized Patients with Type 2 Diabetes Mellitus.” *Diabetes Technology and Therapeutics*, vol. 20, no. 1, 2018, pp. 32–38, doi:10.1089/dia.2017.0250.

Mansfield, Avril, et al. “Use of Accelerometer-Based Feedback of Walking Activity for Appraising Progress With Walking-Related Goals in Inpatient Stroke Rehabilitation.” *Neurorehabilitation and Neural Repair*, vol. 29, no. 9, 2015, pp. 847–57, doi:10.1177/1545968314567968.

Matthews, C. E., et al. “Sources of Variance in Daily Physical Activity Levels as Measured by an Accelerometer.” *Medicine & Science in Sports & Exercise*, vol. 34, no. 8, Aug. 2002, pp. 1376–81, doi:10.1097/00005768-200208000-00021.

Mudge, Suzie, and N. Susan Stott. “Test—Retest Reliability of the StepWatch Activity Monitor Outputs in Individuals with Chronic Stroke.” *Clinical Rehabilitation*, vol. 22, no. 10–11, Oct. 2008, pp. 871–77, doi:10.1177/0269215508092822.

Prajapati, Sanjay K., et al. “A Novel Approach to Ambulatory Monitoring.” *Neurorehabilitation and Neural Repair*, vol. 25, no. 1, 2011, pp. 6–14, doi:10.1177/1545968310374189.

Sanchez, M., et al. “Accelerometric Assessment of Different Dimensions of Natural Walking during the First Year after Stroke: Recovery of Amount, Distribution, Quality and Speed of Walking.” *Journal of Rehabilitation Medicine*, vol. 47, no. 8, Foundation for Rehabilitation Information, Sept. 2015, pp. 714–21, doi:10.2340/16501977-1994.

Urbin, M. A., et al. “Acceleration Metrics Are Responsive to Change in Upper Extremity Function of Stroke Survivors.” *Archives of Physical Medicine and Rehabilitation*, vol. 96, no. 5, W.B. Saunders, May 2015, pp. 854–61, doi:10.1016/j.apmr.2014.11.018.

Uswatte, Gitendra, et al. “Validity of Accelerometry for Monitoring Real-World Arm Activity in Patients With Subacute Stroke: Evidence From the Extremity Constraint-Induced Therapy Evaluation Trial.” *Archives of Physical Medicine and Rehabilitation*, vol. 87, no. 10, Oct. 2006, pp. 1340–45, doi:10.1016/j.apmr.2006.06.006.

Velez, Mariel, et al. “Tracking Generalized Tonic-Clonic Seizures with a Wrist Accelerometer Linked to an Online Database.” *Seizure*, vol. 39, W.B. Saunders Ltd, July 2016, pp. 13–18, doi:10.1016/j.seizure.2016.04.009.

Weeks, Douglas L., et al. “Implementing Wearable Sensors for Continuous Assessment of Daytime Heart Rate Response in Inpatient Rehabilitation.” *Telemedicine and E-Health*, vol. 24, no. 12, Dec. 2018, pp. 1014–20, doi:10.1089/tmj.2017.0306.

Weenk, Mariska, et al. “Continuous Monitoring of Vital Signs Using Wearable Devices on the General Ward: Pilot Study.” *JMIR MHealth and UHealth*, vol. 5, no. 7, 2017, p. e91, doi:10.2196/mhealth.7208.

Weiss, Aner, et al. “Objective Assessment of Fall Risk in Parkinson’s Disease Using a Body-Fixed Sensor Worn for 3 Days.” *PLoS ONE*, edited by Antony Bayer, vol. 9, no. 5, Public Library of Science, May 2014, p. e96675, doi:10.1371/journal.pone.0096675.