Statistical Analysis and Results

# 1. Statistical Analysis

The present analysis was not pre-registered as we had no a priori hypotheses and thus, given the pilot nature of this study, was considered exploratory. Inferential statistics were treated as highly unstable local descriptions of the relations between model assumptions and data in order to acknowledge the inherent uncertainty in drawing generalised inferences from single and small samples (Amrhein, Trafimow, et al., 2019). For all analyses we opted to avoid dichotomising the existence of effects and therefore did not employ traditional null hypothesis significance testing on parameter estimates (Amrhein, Greenland, et al., 2019; McShane et al., 2019). Instead, we opted to take an estimation-based approach instead (Cumming, 2014). For all analyses model parameter estimates and their precision, along with conclusions based upon them, were interpreted continuously and probabilistically, considering data quality, plausibility of effect, and previous literature, all within the context of each model. We focused primarily on qualitative examination of our results based on visualization of the data and models for fixed effects, and exploration of variances using random effects and visualising individual participant level model predictions.

All analysis was performed in R (version 4.2.2, The R Foundation for Statistical Computing, 2022) and all data and code is presented in the supplementary materials (https://osf.io/kf9r3/). The aim of our analysis was to explore how well the Zelemiq sensor device predicted blood lactate levels as measured from capillary samples by the Biosen C-Line. The dependent variable in our model was therefore the blood lactate levels, and the independent predictor variable was the Zelemiq sensor data. A rolling mean was calculated for the Zelemiq data using a 10 sample window prior to the corresponding blood lactate values at that time point. A linear mixed effect model was estimated (Bates et al., 2023) using Restricted Maximal Likelihood with the Zelemiq sensor data as a fixed effect, and allowing random intercepts and slopes by participant id. The model equation was as follows:

We initially fit the model based on the unadjusted averaged Zelemiq sensor data. However, it was clear from this initial model that the model was a poor fit to the data (see posterior predictive check here: https://osf.io/98ey6) likely resulting from the considerable variance in the intercepts between participants due to very different baseline Zelemiq sensor values (see model and individual predictions here: https://osf.io/j26ea). As such, we re-centered within each individual participant based upon their baseline average Zelemiq sensor value (i.e., within each participant ). Comparison of the re-centred model with the original model based on Bayes factors calculated with approximate Bayesian Information Criterion (Wagenmakers, 2007) suggested there was very strong evidence supporting the re-centred model ( 22.8; Kass & Raftery (1995)). Thus we proceeded with this model. The assumption checks for the re-centered model can be seen in the supplementary materials (see https://osf.io/kbzsa). Slope parameters for the Zelemiq, both fixed and random, are divided by 100 to aid interpretability such that they refer to the change in blood lactate for a 0.01 unit change in the Zelemiq sensor value. Profile confidence intervals (CI) were calculated for all model parameters at the 95% level. We examined adjusted and unadjusted intraclass correlation coefficient (ICC) and the marginal and conditional (Nakagawa et al., 2017). Model predictions were visualised for both the conditional fixed effect, and at the individual participant level using the marginaleffects package (Arel-Bundock et al., 2022). The raw Zelemiq sensor data (i.e., unaveraged) was also visualised alongside the blood lactate data. All data visualisations were made using ggplot2 (Wickham et al., 2022), the tidybayes package (Kay & Mastny, 2022), and the patchwork package (Pedersen, 2022).

# 2. Results

Model parameters estimates and confidence intervals are shown in [Table 1](#tbl-model) and [Figure 1](#fig-model) shows the raw data and model predictions for each individual participant in addition to the conditional model predictions. The fixed effect for the intercept corresponded to a blood lactate value of 1.35[95%CI 0.87 to 1.83] and the fixed effect for the slope suggested that blood lactate increased by 0.7[95%CI 0.48 to 1.02] per 0.01 unit increase in the Zelemiq sensor value (see [Figure 1](#fig-model) (C)).

Table 1: Model parameter estimates for both fixed and random effects.

| Model Term | Estimate | Lower 95% CI | Upper 95% CI |
| --- | --- | --- | --- |
| **Fixed Effects** | | | |
| Intercept | 1.35 | 0.87 | 1.83 |
| Zelemiq | 0.70 | 0.48 | 1.02 |
| **Random Effects** | | | |
| *σIntercept* | 0.52 | 0.15 | 1.03 |
| *σZelemiq* | 0.28 | 0.10 | 0.68 |
| *ρIntercept:Zelemiq* | 0.66 | -0.30 | 1.00 |
| *σResidual* | 0.78 | 0.64 | 1.00 |
| Note: |  |  |  |
| CI = confidence interval |  |  |  |
| Slopes for Zelemiq divided by 100 to aid interpretation such that they refer to the change in blood lactate for a 0.01 unit change in the Zelemiq sensor value | | | |

|  |
| --- |
| Figure 1: Panel (A) shows the raw Zelemiq sensor output and Biosen C-Line blood lactate during the course of the incremental test where time has been normalised to 0-100% of the test duration with a locally estimated scatter smooth (LOESS) curve for each, panel (B) shows the individual participant level predicted values (thick lines) with 95% confidence intervals (ribbons) from the fitted model, and panel (C) shows the conditional model predicted values (thick lines) with 95% confidence intervals (ribbons). Individual points in each panel are the individual observed values of data. |

Considering the random effects there was not substantial variance in the individual participant intercept values with a = 0.52[95%CI 0.15 to 1.03] reflecting typical variance in resting blood lactate. However, there was more conspicuous variance in the slopes across individual participants with a = 0.28[95%CI 0.1 to 0.68] . The variation in slopes can be seen clearly in [Figure 1](#fig-model) (B). The correlation between random intercepts and slopes was positive, though the estimate was imprecise ( = 0.66[95%CI -0.3 to 1]) Both the adjusted compared with the unadjusted ICC (adjusted = 0.78, unadjusted = 0.25), and the marginal compared with the condition (marginal = 0.67, conditional = 0.93) suggested that the majority of variance in the model was attributable to the individual participant level.

# 3. References

Amrhein, V., Greenland, S., & McShane, B. (2019). Scientists rise up against statistical significance. *Nature*, *567*(7748), 305–307. <https://doi.org/10.1038/d41586-019-00857-9>

Amrhein, V., Trafimow, D., & Greenland, S. (2019). Inferential Statistics as Descriptive Statistics: There Is No Replication Crisis if We Don’t Expect Replication. *The American Statistician*, *73*(sup1), 262–270. <https://doi.org/10.1080/00031305.2018.1543137>

Arel-Bundock, V., Diniz, M. A., & Greifer, N. (2022). *Marginaleffects: Marginal Effects, Marginal Means, Predictions, and Contrasts*. <https://CRAN.R-project.org/package=marginaleffects>

Bates, D., Maechler, M., Bolker [aut, B., cre, Walker, S., Christensen, R. H. B., Singmann, H., Dai, B., Scheipl, F., Grothendieck, G., Green, P., Fox, J., Bauer, A., & simulate.formula), P. N. K. (shared. copyright on. (2023). *lme4: Linear Mixed-Effects Models using ’Eigen’ and S4*. <https://cran.r-project.org/web/packages/lme4/index.html>

Cumming, G. (2014). The New Statistics: Why and How. *Psychological Science*, *25*(1), 7–29. <https://doi.org/10.1177/0956797613504966>

Kass, R. E., & Raftery, A. E. (1995). Bayes Factors. *Journal of the American Statistical Association*, *90*(430), 773–795. <https://doi.org/10.1080/01621459.1995.10476572>

Kay, M., & Mastny, T. (2022). *Tidybayes: Tidy Data and ’Geoms’ for Bayesian Models*. <https://CRAN.R-project.org/package=tidybayes>

McShane, B. B., Gal, D., Gelman, A., Robert, C., & Tackett, J. L. (2019). Abandon Statistical Significance. *The American Statistician*, *73*(sup1), 235–245. <https://doi.org/10.1080/00031305.2018.1527253>

Nakagawa, S., Johnson, P. C. D., & Schielzeth, H. (2017). The coefficient of determination R2 and intra-class correlation coefficient from generalized linear mixed-effects models revisited and expanded. *Journal of The Royal Society Interface*, *14*(134), 20170213. <https://doi.org/10.1098/rsif.2017.0213>

Pedersen, T. L. (2022). *Patchwork: The Composer of Plots*. <https://CRAN.R-project.org/package=patchwork>

Wagenmakers, E.-J. (2007). A practical solution to the pervasive problems ofp values. *Psychonomic Bulletin & Review*, *14*(5), 779–804. <https://doi.org/10.3758/BF03194105>

Wickham, H., Chang, W., Henry, L., Pedersen, T. L., Takahashi, K., Wilke, C., Woo, K., Yutani, H., Dunnington, D., & RStudio. (2022). *ggplot2: Create Elegant Data Visualisations Using the Grammar of Graphics*. <https://CRAN.R-project.org/package=ggplot2>