University for Continuing Education Krems

Statistical analysis plan

Digital Detox Study

A randomized controlled trial with two unblinded, parallel arms assessing the effect of reducing smartphone screen time on mental health

Authors:	Katja Haider, BA MSc, Univ.-Prof. Dr. Christoph Pieh Assoz. Prof. Priv.-Doz.Dipl.-Ing.Dr. Elke Humer, MSc BEd Bakk.

Version:	01 (original plan)
Clinical trial phase:	n.a.
Release date:	March 19, 2025

Table of contents

1. Purpose of the analyses

This statistical analysis plan (SAP) should provide a detailed overview of the planned strategy for the statistical analysis and methodology of the data from the study protocol "Digital Detox Study: A randomized controlled trial with two unblinded, parallel arms assessing the effect of reducing smartphone screen time on mental health" (version 1).

The purpose of this SAP is to describe the statistical approaches to analyze the collected study data. The SAP is meant to supplement the above-mentioned study protocol.

2. Protocol summary

The aim of the Digital Detox study is to examine the effect of reduced smartphone use on mental health. The Digital Detox study is designed as a randomized controlled trial with two unblinded, parallel arms (intervention group and control group). Participants will be randomly assigned to two groups. Randomization will take place after an initial eligibility screening. The screening takes place when participants enter the study after they have given their informed consent to participate. After the randomization, the intervention group will start the intervention, a three-week digital detox in which smartphone use is to be reduced to ≤ 2 hours per day, while the control group continues to use their smartphone as before. The intervention phase starts on a Monday and ends on a Sunday three weeks later (see Schedule below). On the Sunday, before the start of the intervention, the mental health parameters (depressive symptoms (PHQ-9), sleep quality (ISI), stress (PSQ-20), well-being (WHO-5), loneliness (three-item loneliness scale), craving (CEQ-F), and physical activity) are measured (baseline survey, t0). At the end of the intervention (Sunday three weeks later), these are measured again (post-intervention survey, t1). After the end of the intervention, the intervention group also returns to their normal, unrestricted smartphone use. Three weeks after the end of the intervention (again on a Sunday), a follow-up survey of the mental health parameters (same as for t0 and t1) takes place (follow-up survey, t2). Baseline, post-intervention and follow-up surveys take place in parallel for both groups. Both groups will also complete items on loneliness and physical activity twice during the intervention (after the first and second week of the intervention) to monitor changes in these variables during the intervention. From the beginning of the intervention (t0) until the follow-up time point (t2), the smartphone screen time of the past week, the smartphone activations, and the most frequently used apps are recorded in the study app via self-report and uploading a screenshot every Monday. Monday was chosen to have a complete picture of the past week's smartphone usage behavior. The study will be performed entirely via a smartphone app (ESMira), which is designed especially to run longitudinal studies.

Schedule

			week 1		week 2		week 3		week 4		week 5		week 6		
upon entering		**Sun**	**Mon**	**Sun**	**Mon**	**Sun**	**Mon**	**Sun**	**Mon**	**Sun**	**Mon**	**Sun**	**Mon**	**Sun**	**Mon**
	CG	no intervention	no intervention		no intervention		no intervention		no intervention		no intervention		no intervention		
	IG	no intervention	Detox (≤2h)		Detox (≤2h)		Detox (≤2h)		no intervention		no intervention		no intervention		
Elibility screening: Sociodemographic variables Smartphone screen time ≥ 3h/day Psychological, psychotherapeutic, or psychopharmacological treatment PHQ-9 WHO-5 ISI PSQ-20 Loneliness Physical activity	Randomization	**Baseline (t0):** PHQ-9 WHO-5 ISI PSQ-20 CEQ-F Loneliness Physical activity	ST W0 (B)	Weekly (1): Loneliness Physical activity	ST W1	Weekly (2): Loneliness Physical activity	ST W2	**Post-Intervention (t1):** PHQ-9 WHO-5 ISI PSQ-20 CEQ-F Loneliness Physical activity	ST W3		ST W4		ST W5	**Follow-up (t2):** PHQ-9 WHO-5 ISI PSQ-20 CEQ-F Loneliness Physical activity	ST W6

Note. CG = control group; IG = intervention group; ST = screen time; Detox (≤2h) = smartphone screen time restriction to a maximum of 2 hours per day; Sun = Sunday; Mon = Monday; W = week, PHQ-9 = Patient Health Questionnaire; WHO-5 = World Health Organization Well-Being Index; ISI = Insomnia Severity Index; PSQ-20 = short form of the Perceived Stress Questionnaire; CEQ-F = Craving Experience Questionnaire.

3. General analysis and reporting conventions

This chapter discusses general policies that will be employed in analyzing and reporting the study data. Deviations from these general policies are necessary, these will be described in the respective section of this SAP.

Unless otherwise specified, the following groups will be mentioned in the statistics:

- Intervention group (IG)
- Control group (CG)

SPSS (version 29.0 or higher) or R will be the software used for the analyses of the data.

3.1. Categorial variables

Descriptive statistics will include the number and percentage of subjects in each category. Unless otherwise specified, percentages will be based on the total number of subjects in the reported population. All percentages will be rounded to one decimal place. The number and percentage of subjects will always be presented in the format XX (XX.X%), with the percentage in parentheses. To ensure completeness, all summaries for categorical and discrete variables will include all categories, even if no subjects had a response in a particular category.

3.2. Continuous variables

Summary statistics will include the number of subjects with available data, as well as the mean and standard deviation (SD). The summary statistic *n* will represent the number of subjects. Means will be reported with two decimal places. Standard deviations will be reported with three decimal places.

3.3. Tests of group differences

The associated *p*-values for the tests of group differences will be presented rounded to three decimal places. *P*-Values that are rounded to 0.000 will be reported as "< .001". Significance tests will be performed two-tailed (α = 5%). Test statistics (e.g. t or χ^2) will be reported rounded to three decimal places with their respective degrees of freedom (df) in parentheses. The statistics will be presented in the following format: e.g. t(df) or χ^2(df).

3.4. Repeated-measures analysis of variance

The outcomes of the repeated measures analysis variance (rm-ANOVA) will include the means and SD for each group and time point (t0, t1, t2). Means and SD will be presented as stated in 3.2. The test statistics will be reported as follows: The F-statistic will be presented rounded to three decimal places. The df for the effect (e.g. time) and the error term (residual variance) will be shown in parentheses behind it. The same principles apply to the resulting *p*-value as stated under 3.3. Thus, the test statistics will be presented as follows: *F*(df effect, df error term) = X.XXX; *p* = .XXX. The effect size will be reported as partial eta squared (partial η^2) rounded to three decimal places.

In case of significant rm-ANOVA results, post-hoc tests are performed to compare groups. These are corrected for multiple comparisons. Results will be presented as mean differences, standard error, 95% confidence intervals, and p-values.

3.5. Assessment time points

Assessments will take place at the following time points:

- Eligibility screening: right upon entering the study
- Baseline (t0): Sunday before the start of the intervention period
- Post-intervention (t1): on a Sunday at the end of the intervention period, three weeks after t0
- Follow-up (t2): on a Sunday, three weeks after t1

If assessments after the baseline assessment are missing, missing values will be dealt with via maximum likelihood estimation.

4. Analysis samples

The study sample consists of only one population, which will be primarily analyzed according to intention-to-treat (ITT), i.e. the data of all participants who are successfully screened for eligibility, randomized, and provided baseline data will be analyzed according to their arm allocation regardless of whether they have reached their screen time goal or not.

Depending on the nature of the study data, secondary subgroup analyses of those who strictly adhered to their screen time limit in all three weeks (per protocol) are possible too.

5. Study subjects

Study subjects are the participants taking part in the Digital Detox study.

5.1. Disposition of subjects

The number of participants screened, enrolled, excluded, randomized/allocated as well as the number of those who filled in the questionnaires at t0, t1, and t2 or discontinued will be reported.

5.2. Demographics and other baseline characteristics

The following variables are collected in the study:
Assessed only in the screening:

- Sociodemographic variables: Age (in years), gender, country of residence, highest completed education, current educational/employment status, urban/rural classification of residence.
- Mental health (self-report):
    - Do you have a mental health condition for which you are currently receiving treatment? (yes/no)
    - Are you currently taking medication for a mental health condition? (yes/no)
    - Smartphone usage: Do you use your smartphone for an average of ≥3 hours per day? (yes/no)

Assessed in screening and at all subsequent time points (t0, t1, t2):

- Depressive Symptoms (PHQ-9): The Patient Health Questionnaire (PHQ-9) is used to assess depressive symptoms in its validated German version, consisting of 9 items (Kroenke et al., 2001; Löwe et al., 2004).
- Well-being (WHO-5): The validated German version of the World Health Organization Well-Being Index (WHO-5) is used to assess participants' well-being and consists of 5 items (Brähler et al., 2007; Topp et al., 2015).
- Sleep Quality (ISI): The Insomnia Severity Index (ISI) is used in its validated German version to assess the quality of sleep and consists of 7 items (Bastien et al., 2001; Gerber et al., 2016; Morin et al., 2001).
- Stress (PSQ-20): Stress is assessed using the German version of the short form of the Perceived Stress Questionnaire (PSQ-20), which consists of 20 items (Fliege et al., 2009; Fliege et al.; 2005; Levenstein et al., 1993).
- Loneliness (3-Item Loneliness Questionnaire): Loneliness is assessed using the German version of the Three Items Loneliness Scale, consisting of 3 items (Klein et al., 2021; Reinwarth et al., 2024).
- Physical activity: "On how many of the last 7 days were you physically active for at least 60 minutes?" and "How many minutes did you spend doing sport in the last 7 days?"
- Craving (CEQ-F) *[not in screening]*: Craving is measured using the Craving Experience Questionnaire (CEQ-F) in the German version, for behavioral addictions and smartphone use in particular, and consists of 9 items (Brandtner & Wegmann, 2023; Cornil et al., 2019; Fritz, n.d.).
- Smartphone use (self-report and upload of screenshots of screen time over the last week, activations, and the most frequently used apps).

5.3. Prior and concomitant medication

During the screening, it is assessed whether participants are currently taking psychopharmacological medication to treat a mental illness. If so, this meets the exclusion criteria for the study and that participant cannot take part in the study. Other medications are not assessed.

5.4. Medical history

During the screening, information is collected not only about the use of psychopharmacological medication but also about whether the participants are currently undergoing psychological or psychotherapeutic treatment. If so, this also meets the exclusion criteria of the study, and that participant cannot take part in the study. In addition, questionnaires on mental health are to be completed (see 5.2). However, this does not result in any exclusion criteria. Other factors of the participants' medical history will not be collected.

6. Study operations

6.1. Protocol deviations

Any non-compliance with the study protocol is considered a protocol deviation. Deviations from the protocol will be avoided. However, if they occur, they will be documented and a reason for the deviation will be given.

6.2. Randomization

After the initial eligibility screening, participants will be randomly assigned to one of the two study groups (IG or CG). A simple randomization based on chance is carried out with SPSS. First, each user-id generated by ESMira (for each participant) is assigned a random number between 0-1. Then, based on these random numbers, the participants are divided into those above and below 0.5, and the values 1 (IG) and 2 (CG) are assigned for each group. The following syntax is used for this:

```
COMPUTE rand=RV.UNIFORM(0,1).
EXECUTE.

COMPUTE group = 1.
IF (rand >= 0.5) group = 2.
VALUE LABELS group 1 'IG' 2'CG'.
EXECUTE.
```

The groups should be roughly the same size, but an exact balance is not guaranteed due to chance.

6.3. Measures of treatment compliance

The planned study does not speak of treatment but of intervention. Compliance with the intervention is given if a participant in the intervention group does not exceed the daily screen time maximum of 2 hours. If a participant has not adhered to the limit, this is referred to as non-compliance. This can be monitored by the study team on the basis of the information on screen time (self-report and screenshot) and is noted as an adherence variable for each intervention week in the data set.

7. Endpoint evaluation

7.1. Overview of analysis methods

7.1.1. Multicenter studies

Not applicable. This is a single-center study.

7.1.2. Timing of analyses

All final analyses will be performed after the last smartphone usage assessment after t2 and all data have been processed and integrated into the study database.

7.1.3. Multiple comparisons

Unless otherwise indicated, all statistical tests will be performed at a two-sided significance level of 5%. The post hoc pairwise comparisons will be adjusted for multiplicity.

7.2. Primary endpoints

The primary endpoints of this study are the mental health assessments for depressive symptoms (PHQ-9), stress (PSQ-20), well-being (WHO-5), and sleep quality (ISI). Cronbach's alphas (internal consistency) are also calculated for each scale at each point in time.

7.2.1. Computation of the primary endpoints

The sum scores of the scales of the primary endpoints (mentioned under 7.2) will calculated according to the analysis rules of the respective scale. No data at t0 is considered a discontinuation (participant withdrawal) and will be excluded from the analyses. Missing data at t1 and/or t2 will be dealt with via maximum likelihood estimation.

7.2.2. Primary analysis of the primary endpoints

In order to examine the effect of smartphone screen time reduction on the participants' mental health (see 7.2), repeated measures ANOVAs (rm-ANOVAs) will be performed. The primary hypotheses focus on the difference between the baseline and the post-intervention time points (t0 – t1) in the intervention group as well as on the difference between the control and the intervention group at t1. An intention-to-treat analysis will be carried out, i.e. participants are evaluated according to their group allocation, regardless of whether they have strictly adhered to the screen time limit or not. Arm allocation (control group or intervention group) will be the between-group factor. Time (t0, t1) will be the within-group factor. Effect sizes will be presented as partial eta squared (η^2 = time × group) or group mean differences. If there are group differences in one of the assessed variables (sociodemographic or mental health) at baseline, this variable can be included as a covariate in the model (e.g. significant group difference in depressive symptoms for the PHQ-9 model). Group differences are calculated using χ^2-Test or independent samples t-tests, provided that all test requirements are met. Non-parametric tests can be used as an alternative.

Percentage changes in the mental health outcomes between the time points will be calculated (Hurst & Bolton, 2004).

The advantage of the ITT analyses is that they reflect the real-world conditions of the intervention and also take into account factors such as a lack of compliance. However, the effectiveness of the intervention may be underestimated because people who did not adhere to the screen time limit are also included. For this reason, secondary analyses are also planned:

7.2.3. Secondary analysis of the primary endpoints

In a secondary analysis, the difference in the primary outcomes between the baseline (t0), the post-intervention (t1), and the follow-up time points (t2) in the intervention group as well as the difference between the control and the intervention group at t2 in the primary outcomes will be additionally examined using the same statistical approach as described above.

Furthermore, subgroup analyses are planned, especially of the subgroup from the intervention group that strictly adhered to the screentime limit (per-protocol), but also with regard to sociodemographic variables (e.g., gender or age) or pre-intervention (t0) mental health parameter scores.

Per-protocol analyses show the maximum effect of the intervention under optimal conditions, but have the disadvantage that practical applicability may be biased, since only "disciplined" participants are considered and the reality is thus not fully reflected.

7.3. Secondary endpoints

The secondary endpoints of this study are the variables for craving (CEQ-F), loneliness, and physical activity. Cronbach's alphas (internal consistency) are also calculated for each scale at each point in time.

7.3.1. Computation of the secondary endpoints

The sum scores of the scales of the secondary endpoints (mentioned under 7.3) will calculated according to the analysis rules of the respective scale. Missing data at t1 and/or t2 will be dealt with via maximum likelihood estimation.

7.3.2. Primary and secondary analysis of the secondary endpoints

The secondary endpoints will be analyzed according to the example of the primary endpoints (described under 7.2.2 and 7.2.3).

7.4. Examination of subgroups

The analysis of subgroups is described in 7.2.3 and 7.3.2.

8. Safety evaluation

Not applicable. No clinical study.

9. Pharmacokinetic evaluation

Not applicable.

10. Interim analysis and data monitoring

Not applicable.

11. References

Bastien, C. H., Vallières, A., & Morin, C. M. (2001). Validation of the Insomnia Severity Index as an outcome measure for insomnia research. *Sleep Medicine, 2*(4), 297-307. https://doi.org/10.1016/s1389-9457(00)00065-4

Brähler, E., Mühlan, H., Albani, C., & Schmidt, S. (2007). Teststatistische Prüfung und Normierung der deutschen Versionen des EUROHIS-QOL Lebensqualität-Index und des WHO-5 Wohlbefindens-Index. *Diagnostica, 53*(2), 83-96. https://doi.org/10.1026/0012-1924.53.2.83

Brandtner, A., & Wegmann, E. (2023). The fear in desire: linking desire thinking and fear of missing out in the social media context. *BMC Psychology, 11*(1), 176.

Cornil, A., Long, J., Rothen, S., Perales, J. C., de Timary, P., & Billieux, J. (2019). The gambling craving experience questionnaire: Psychometric properties of a new scale based on the elaborated intrusion theory of desire. *Addictive Behaviors, 95*, 110-117.

Fliege, H., Rose, M., Arck, P., Levenstein, S., & Klapp, B. F. (2009). PSQ – Perceived Stress Questionnaire. In. Leibniz Institute for Psychology – Open Test Archive. https://doi.org/https://doi.org/10.23668/psycharchives.5138

Fliege, H., Rose, M., Arck, P., Walter, O. B., Kocalevent, R.-D., Weber, C., & Klapp, B. F. (2005). The Perceived Stress Questionnaire (PSQ) Reconsidered: Validation and reference values from different clinical and healthy adult samples. *Psychosomatic Medicine, 67*(1). https://doi.org/10.1097/01.psy.0000151491.80178.78

Fritz, M. The Craving Experience Questionnaire: Validation of a German Version. Gerber, M., Lang, C., Lemola, S., Colledge, F., Kalak, N., Holsboer-Trachsler, E., Pühse, U., & Brand, S. (2016). Validation of the German version of the Insomnia Severity Index in adolescents, young adults and adult workers: results from three crosssectional studies. *BMC Psychiatry, 16*, 174. https://doi.org/10.1186/s12888-016-0876-8

Hurst, H., & Bolton, J. (2004). Assessing the clinical significance of change scores recorded on subjective outcome measures. *Journal of Manipulative and Physiological Therapeutics*, *27*(1), 26–35. https://doi.org/10.1016/j.jmpt.2003.11.003

Klein, E. M., Zenger, M., Tibubos, A. N., Ernst, M., Reiner, I., Schmalbach, B., ... & Beutel, M. E. (2021). Loneliness and its relation to mental health in the general population: Validation and norm values of a brief measure. *Journal of Affective Disorders Reports, 4*, 100120.

Kroenke, K., Spitzer, R. L., & Williams, J. B. (2001). The PHQ-9: validity of a brief depression severity measure. *Journal of General Internal Medicine, 16*(9), 606-613. https://doi.org/10.1046/j.1525-1497.2001.016009606.x

Levenstein, S., Prantera, C., Varvo, V., Scribano, M. L., Berto, E., Luzi, C., & Andreoli, A. (1993). Development of the perceived stress questionnaire: A new tool for psychosomatic research. *Journal of Psychosomatic Research, 37*(1), 19-32. https://doi.org/https://doi.org/10.1016/0022-3999(93)90120-5

Löwe, B., Spitzer, R. L., Gräfe, K., Kroenke, K., Quenter, A., Zipfel, S., Buchholz, C., Witte, S., & Herzog, W. (2004). Comparative validity of three screening questionnaires for DSM-IV depressive disorders and

physicians' diagnoses. *Journal of Affective Disorders, 78*(2), 131-140. https://doi.org/10.1016/S0165-0327(02)00237-9

Morin, C. M., Belleville, G., Bélanger, L., & Ivers, H. (2011). The Insomnia Severity Index: psychometric indicators to detect insomnia cases and evaluate treatment response. *Sleep, 34*(5), 601-608. https://doi.org/10.1093/sleep/34.5.601

Pieh, C., Humer, E., Hoenigl, A., Schwab, J., Mayerhofer, D., Dale, R., & Haider, K. (2025). Smartphone screen time reduction improves mental health: a randomized controlled trial. *BMC Medicine*, *23*(1), 107. https://doi.org/10.1186/s12916-025-03944-z

Reinwarth, A. C., Petersen, J., Beutel, M. E., Hautzinger, M., & Brähler, E. (2024). Social support in older adults: Validation and norm values of a brief form of the Perceived Social Support Questionnaire (F-SozU K-6). *Plos one, 19*(3), e0299467.

Topp, C. W., Østergaard, S. D., Søndergaard, S., & Bech, P. (2015). The WHO-5 Well- Being Index: A Systematic review of the literature. *Psychotherapy and Psychosomatics, 84*(3), 167-176. https://doi.org/10.1159/000376585

12. List of abbreviations

χ^2	Chi squared
CG	control group
df	degrees of freedom
e.g.	for example
F	F-statistic
i.e.	id est, that is
IG	intervention group
ITT	intention-to-treat
partial η^2	partial eta squared
rm-ANOVA	repeated measures analysis of variance
SAP	statistical analysis plan
SD	standard deviation
t	t-statistic