LIVE@Home.Path – a mixed-method, stepped-wedge, randomized controlled trial to assess a clinical pathway for home-dwelling people with dementia and their caregivers

Statistical Analysis Plan (SAP)

Version 17.01.2023

NFR Project number 273581

Introduction

Preface

The objective of the Statistical Analysis Plan (SAP) is to outline the statistical methodology to be used for the final analysis of LIVE@Home.Path[1].

The SAP is based on the following documents

Study Protocol	Version 4.0 (26.04.2019)
Amendments to Study Protocol / Observational Plan	Change in protocol due to Covid-19 restrictions. Approval for change from Regional Ethical Committee (REK: 10861, 06.04.2020), update clinicaltrials.gov (31.08.2020)
Scientific Review Committee Charter (SRC)	22.11.2017 Research Council Norway granted the LIVE@Home.Path allocation 19.06.2018 Dato for signing the contract 06.05.2019 Approval Ethic Committee (2019/385/REK North)
Professional development Plan (PD)	Post-doc 1: 20210311 Post-doc 2: 20211209 Post-doc 3: 20201209
Statistical Analysis Plan for the Interim Analyses	Not applicable

Timing of statistical analyses

Final analysis after the last data collection period is completed.

Modification History

Changes to the study protocol

Protocol version 4.0 was changed in April 2020 due to the Covid-19 restrictions, as the pandemic severely hampered the implementation of the intervention which had to be delivered at the participants homes. We therefore had to change the design of the overall trial, postponing the intervention in two municipalities and delivering the intervention by phone in the last municipality. In addition, we initiated the PAN.DEM cohort in the LIVE@Home.Path trial, collecting data from phone interviews with caregivers on change in health service use and neuropsychiatric symptoms, risk perception and restrictions [2]. Changes in design approved by Ethical committee (REK: 10861), update on clinicaltrials.gov (31.08.2020)

Study Design

Indication	
Design	Stepped-wedge design

	Closed cohort where individuals are identified at the start of the trial and scheduled for repeated outcome assessment [3]
Phase	
Primary objective	To determine treatment effect on caregiver burden measured as Resource Utilization Dementia (RUD), and the Relative Stress Scale (RSS). Based on the correlation coefficient, we will use all three subcategories of RUD separately. Total score will be used when assessing RSS. Further, we will evaluate the economic consequences of the treatment by calculating Incremental Cost-Effectiveness Ratio (ICER) and quality adjusted life years (QALY).
Secondary objectives	To determine treatment effect on quality of life (QoL) using the European Quality of Life – 5 dimension – 5 levels (EQ-5D-5L), EQ-5D-VAS scale, Quality of Life in Alzheimer's disease scale (QoL-AD); neuropsychiatric symptoms in dementia using the Neuropsychiatric Inventory 12 item (NPI-12), agitation in dementia using Cohen-Mansfield Agitation Inventory (CMAI), depression in old age (Geriatric Depression Scale - GDS), functional level for instrumental and personal activities (ADL and IADL), and the caregiver rating of clinical meaningful change using Clinical Global Impression of Change (CGIC).
Exploratory objectives	In line with total score of RSS, we will also explore the treatment effect on both single and clustered items. In line with total score of QoL based instruments (EQ-5D-5L, QoL-AD, QoL – CG, GDS), we will also explore the treatment effect on both single and clustered items To assess the effect on pain in dementia using the Mobilization – Observation – Behavior – Intensity Dementia Pain Scale (MOBID-2) To determine the characteristics of the dyads who had the highest improvement in RUD and RSS, i.e., those who benefitted the most from the intervention.
Endpoints	Primary endpoint: Change in caregiver burden and informal care time among the caregivers in the intervention group during active intervention period, and the follow-up. Change in Cost-efficiency of treatment. Secondary endpoint: Improvement in QoL for both PwD and caregivers. Reduction in neuropsychiatric symptoms in dementia and functional level during active intervention period and follow-up. Exploratory endpoint: Reduction in pain during active intervention period.
Treatments	Allocation to municipal coordinator delivering the multicomponent LIVE intervention (L:learning, I: innovation, V: volunteering and E:empowerment) or care as usual.
Number of patients	315 dyads will be screened to achieve at least 105 dyads per municipality.

Safety analysis	Not applicable
Planned enrolment	3 municipalities and 9 municipality districts

Sample size estimation

The required sample size calculated before recruitment of participants aimed to detect a difference of 7 h/week for the primary outcome RUD. Based on the literature[4], we assumed that the number of hours of informal care is 46 h/week with a standard deviation (SD) of 20 h/week. With 80% power and a significance level of 5%, the required sample size was estimated to be 260 dyads. To allow for 20% loss to follow-up, a total of 315 dyads, equaling 105 per municipality, must be included.
The post-hoc sample size calculations based on the outcome PADL showed that by treating the study as an individually randomized stepped wedge design[3], with 3 intervention sequences, 5 periods, effect size of 7h/week, outcome variance of 400, significance level of 5% and within-individual correlation of 0.1075 over the 24 month trial would require a total of 240 dyads to be included. To allow for 20% loss to follow-up, the sample size must reach 288 dyads.

Randomization, blinding and unblinding procedures

Block randomization was performed using a computerized random number generator to allocate individual dyads to the three intervention sequences. We will apply stratification to secure that about one third of the dyads in each treatment sequence will be recruited from each municipality. Research assistants, researchers conducting the analyses and other study personal conducting data collection will be blinded to the randomization order, the dyads will be blinded to group allocation until contacted by their coordinator at start of the intervention period. Given the practice change of the intervention, the municipality home care services and coordinators will be informed about allocation sequence when their clusters entered the intervention period.

Analysis Sets

Full Analysis Set (FAS)

The intention-to-treat population includes all randomized dyads who were assigned a coordinator by the municipality. The dyads will be analyzed separately (i.e., PwD and caregiver), dependent on outcome of interest.

This will be the primary analysis set for baseline summary statistics of demographics, primary, secondary and exploratory efficacy analysis as elaborated in section 10.

Municipal coordinator set (MCS)

All coordinators in all municipality districts

Usage overview of analysis sets table

Analyses	FAS	MCS	
Disposition			
Demographics and background	x		
Medical history			
Efficacy: primary	x		
Efficacy: secondary	x		x
Efficacy: other	x	x	x

General Statistical Methods and Definitions

General statistical methods

The statistical analysis will be presented by treatment group for the different analysis sets as outlined in section 4.

Stepped wedge design (SWD) do not strictly feature a treatment and control group, and thus, the summary statistics will include all participants at baseline, and the characteristics according to intervention sequences.

Continuous variables will be summarized using descriptive statistics (means and interquartile ranges), and categorical variables will be presented by frequencies and percentages. A table will present the number of caregivers and PwD with the following data: missing values, mean, median, standard deviation and confidence intervals.

Means and median as well as percentages will be presented with 1 additional decimal place and standard deviation by 2 additional decimals.

If the number of PwD or caregivers in a category is 0, then percentage will not be displayed.

P-values will be reported to 3 decimal places.

Unless otherwise specified, all statistical tests will be two-sided.

Causal model

Due to the longitudinal structure of the collected trial data, we considered the following linear mixed model to estimate and identify the treatment effect due to the intervention. Our SWD, consisting of $j = 3$ municipalities and 3 steps, is a closed cohort design, and thus, time related confounders are important to adjust for [5] when assessing treatment effect.

Mixed-effects regression is one of the most common approaches when estimating the intervention effect in SWD [6, 7]. In our design, we will consider the following model as our primary analysis model, assuming a constant treatment effect.

(1) $Y_{ijt} = \beta_0 + O_{ijt}\beta_1 + X_{ijt}\beta_2 + b_i + \epsilon_{ijt}$

where notation Y_{ijt} refers to the primary outcome caregiver burden (RUD and RSS), elaborated in section 3 Study Design, for individual i nested in municipality j at time point t. O_{ijt} is the treatment indicator (effect of receiving assistance from a municipality coordinator) and β_1 is the average treatment effect parameter of interest. Since the study collects repeated observations from the same participants, we consider a longitudinal regression model that adjusts for participant-level random effect, $b_i \sim N(0, \sigma_b^2)$, which corresponds to a compound symmetric correlation structure.

$\epsilon_{ijt} \sim N(0, \sigma_\epsilon^2)$ represents random error that changes across t and i.

X_{ijt} constitutes control variables which are demographic, socio-economic and clinical confounders (age, sex, caregiver marital status, work status, ADL function), and stratification variables used in the design stage.

Due to the non-normal distribution of hours and costs we will use both linear mixed-effects, as well as general lineal mixed regression model for assessing both primary variables. For RSS, and other continuous variables, we use Linear Mixed regression model. Log-transformation of the outcome will also be considered to address potential non-normality in the response.

Intervention effect

In secondary analysis, we will further examine the extent to which the treatment effect will depend on the calendar time when COVID lockdown happened, and the potential for delayed effect. We collectively refer these effects as heterogeneous treatment effect, and provide a schematic illustration following Section 3.2 of Li et al. (2020) [6]. The rationale is that there the lockdown during the pandemic may contribute to change in the intervention effect. In addition, the intervention may be most effective when it is first introduced, and the sustained effect may be smaller. Therefore, when modelling the intervention effect, we will allow for heterogenous effect and assess the robustness of our primary analysis. The table below shows how the intervention effect can differ between periods (t) across different groups (g). The deltas (δ) constitute the intervention effect, and their subscript defines heterogenous effects where δ_0 is the pre-lockdown intervention effect, δ_0^* is the intervention effect post-lockdown, and $\delta_1^*$is the follow-up post-lockdown treatment effect.

When modelling the intervention effect, we will test the null-hypothesis (H_0)

$$\delta_0 = \delta_0^* = \delta_1^*$$

	$t = 1$	$t = 2$	$t = 3$	$t = 4$
$g = 3$			δ_0^*	δ_1^*
$g = 2$		δ_0^*	δ_1^*	δ_1^*
$g = 1$	δ_0	δ_1^*	δ_1^*	δ_1^*

Economic evaluation

If we identify an average treatment effect on one of the primary and secondary outcomes, we will use cost-efficiency (CEA) and cost-utility analysis (CUA) to evaluate the economic effects of the LIVE treatment.

In cost-effectiveness analysis the effect is measured in health units (or other outcome measures depending on the context). The outcomes are common health measures in natural units, and the results are presented as cost per effect unit/outcome measure. In cost-utility analysis, also called cost per QALY analysis, the effect is measured as quality-adjusted life years (QALYs) on a scale of 0 (death) and 1 (full health). A quality-adjusted life year (QALY) captures both life years gained and improved health as a result of treatment [8, 9].

The economic evaluation will be measured at society level.

Costs and QALYs will be measured for each phase of 6 months

The cost-effective analysis will relate the difference in total mean costs at societal level per 6 months between all intervention -and usual care phases to the difference in RUD, RSS, QoLs, QALYs, NPI, function level such as ADL and IADL

The analysis will use a discounting rate at 3.5%.

Following other costs studies calculating RUD in an advanced economy setting, our analysis will use 35% of the Norwegian average income to value lost leisure time for retired caregiver.

The primary economic evaluation will use the opportunity cost method to assess the costs of informal caregiving. For a sensitivity analysis, the replacement cost will be used.

Unit costs will be based on information from Statistics Norway and information from the primary health care institutions in each municipalities.

Covariates and strata

In the main statistical analyses, the following covariates are used for both caregivers and PwD.

For PwD: Age, sex, marital status, cohabitation status, number of co-caregivers, clinical characteristics of dementia (dementia etiology, neuropsychiatric symptoms, functional level)

For Caregivers: Age, sex, marital status, cohabitation status, numbers of co-caregivers involved in caregiving for the PwD, work status, use of health care services,

Groups, municipalities

Municipalities will be grouped in the main analysis and used as an adjustment variable (fixed effects). This is not the case for district municipalities as we consider that the dyads are similar within municipalities

Subgroups

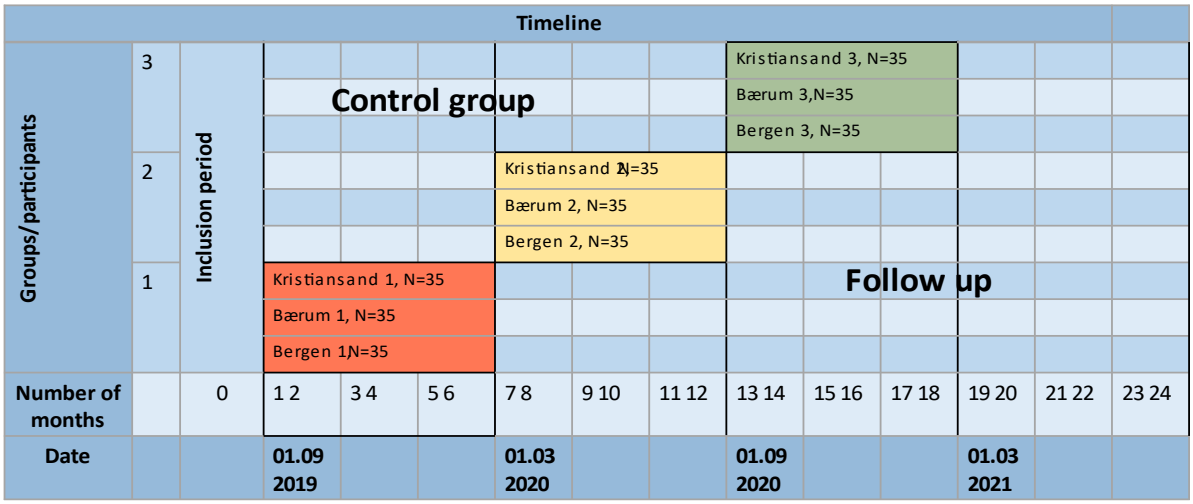
In subgroup analyses, the following covariates will be used to define relevant subgroups:

- Gender (male, female)
- Work (retired and working caregivers)
- Caregiver relation (spouse, children)
- Residency (living alone, or with caregiver)
- Formal caregiver (whether dyads are receiving care by home nursing)
- Categories of caregiver burden intensity
- Categories of dementia diagnoses

Results from subgroups will be presented for the primary analyses.

Observation and analysis times

Study periods



	STUDY PERIOD							
	Enrollment	Allocation	Post-allocation					Close-out
TIMEPOINT	-t	0	t1	t2	t3	t4	t5	
ENROLLMENT:								
Eligibilityscreen	x							
Informed consent	x							
Allocation		x						
INTERVENTIONS:								
1st group			x	x				
2nd group				x	x			
3rd group					x	x		
ASSESSMENTS:								
Baseline variables			x					
Outcome variables				x	x	x	x	x

The stepped wedge randomized control design. A) The randomization in time takes place at month 0. First group (red) is in the intervention period from month 1 to 6, second group (yellow) from month 7 to 12, third group (green) from month 13 to 18. Implementation seminars will be held at month 0, 6 and 12, midway evaluation at month 3,9 and 15. Data will be collected at baseline (month 0), after first intervention period (month 6-7), after second intervention period (month 12-13), after third intervention period (month 18-19) and at end of study at 24 months. B) Schedule of enrollment, interventions and assessments over the study period.

Demographics and Background Characteristics

Demographics and baseline characteristics

The following demographic characteristics will be presented descriptively and listed for both PwD and caregiver:

- Sex (male, female)
- Age in years
- Marital status (spouse, child)
- Work status (employed, unemployed, retired)
- Number of co-caregivers
- MMSE total score
- RUD (number of care hours the last 30 days)
- RSS total score
- NPI total score
- QoL total score

Date of informed consent and protocol version will be listed

Medical history

We will provide information on the number of years since formal diagnosis of dementia together with descriptive statistics.

Exposure and Compliance

Treatment groups

The definition of treatment assignment as specified in section "Analysis sets" will be used in this section, too.

Deviations from the randomized treatment sequence occurred under the first Covid-19 lockdown, se section Modification History, Changes to the study protocol

Treatment duration

Each municipal coordinators are writing logs, and we will collect the data and calculate total hours of care by the coordinator, divided into different categories.

Descriptive statistics for the total hours of care by the coordinators will be presented.

Delivery of the intervention

One coordinator from the municipal per dyad

Efficacy

Primary efficacy analysis

The primary efficacy analysis variable caregiver burden is based on two variables. The first is RUD, divided in three categories: number of hours the last 30 days invested in personal, instrumental and supervisory care. The second is RSS measured as total score but also

specific and clustered items. Lower hours of informal care as well as lower score of RSS, indicates reduction in caregiver burden.

The null hypothesis states that the are no differences in burden and resource utilization between the LIVE intervention and usual care. The primary alternative hypothesis (H1) states that the LIVE intervention will improve caregiver burden and resource utilization compared with usual care. The secondary alternative hypothesis (H2) is that the LIVE intervention will provide more cost-efficient services than usual care.

H0: LIVE treatment = control

H1: LIVE treatment > control

H2: cost efficiency > control

where LIVE treatment denotes the proportion of caregivers experiencing improvement in care burden. The primary hypothesis will be statistically assessed by means of a $\chi 2$ test on the one-sided significance level of 2.5% (corresponding to 5% two-sided).

We will plot the regression line to visualize the parallel effect for both treatment group and control group.

The significance level will be set to 2.5% (one-sided, corresponding to 5% two-sided), the corresponding level of one-sided CI is 97.5% (corresponding to 95% two-sided).

Missing data

We will pre-specify a list of baseline covariates that are likely associated with the missingness and the outcome, and adjust for these covariates in our primary analyses. This covariate-adjusted analysis will be valid under the missing at random (MAR) mechanism.

Sensitivity and robustness tests of regression analyses and economic evaluation

Due to the complexity of setting up a causal model in a stepped-wedge design trial, we will run several robustness and sensitivity analyses for primary and secondary efficacy analyses.

Subgroup analyses (demographic and clinical) will be performed and compared with the primary efficacy analysis

Robustness checks will also be performed using clustered or single items on measurements such as NPI, RSS, GDS and QoL.

In the economic evaluation, we will set up a simulation model where we can test different cost levels. We will plot the resulting ICERs graphically as a ratio between costs and the effectiveness/utility or as a distribution with uncertainty in cost / effectiveness plane. Four quadrants represent all combinations of possible outcomes. The more effective outcomes are located right on the x-axis, and with the rise of y-axis the cost of the outcome rises. An ICER

that is more costly and more efficient than care as usual is located in the first quadrant; in case of a more costly and less efficient treatment, ICER is in second quadrant.

Secondary efficacy analyses

Other efficacy analyses will be based on the secondary objectives (QoL measurements and NPI) for both PwDs and caregivers They will be analyzed on the full analyses set (FAS). Higher scores of QoL (for both EQ-5D-5L and QoL Alzheimer's disease scale) indicates low burden for both PwDs and caregivers. Further, high total scores of NPI indicate high burden for both PwDs and caregivers.

Interim analysis

An interim analysis is not planned.

Software

If not stated otherwise, the data will be analyzed using Stata version 17

Abbreviations

CG – Caregiver
FAS – Full Analysis Set
PwD – Person with dementia
RUD – Resource Utilization Dementia
SD – Standard Deviation
SWD – Stepped Wedge Design
QoL – Quality of Life
MI – Multiple Imputation
MAR – Missing at Random
MNAR – Missing not at Random

References

1. Husebo, B.S., et al., *LIVE@Home.Path-innovating the clinical pathway for home-dwelling people with dementia and their caregivers: study protocol for a mixed-method, stepped-wedge, randomized controlled trial.* Trials, 2020. **21**(1): p. 510.
2. Gedde, M., Husebo, BS, Erdal A, Puaschitz NG, Vislapuu M, Angeles RC, Berge LI., *Acess to and interest in assistive technology for home-dwelling people with dementia during the COVID-19 pandemic (PAN.DEM).* International Review of Psychiatry, 2021.
3. Hooper, R. and C. Knowles, *Improving the efficiency of individually randomized clinical trials by staggering the introduction of the intervention.* Statistics in Medicine, 2019. **38**(1): p. 44-52.
4. Schulz, R., et al., *End-of-life care and the effects of bereavement on family caregivers of persons with dementia.* New England Journal of Medicine, 2003. **349**(20): p. 1936-1942.
5. Hemming, K. and M. Taljaard, *Reflection on modern methods: when is a stepped-wedge cluster randomized trial a good study design choice?* International Journal of Epidemiology, 2020. **49**(3): p. 1043-1052.

6. Li, F., et al., *Mixed-effects models for the design and analysis of stepped wedge cluster randomized trials: An overview.* Statistical Methods in Medical Research, 2020. **30**(2): p. 612-639.
7. Hemming, K., et al., *The stepped wedge cluster randomised trial: rationale, design, analysis, and reporting.* BMJ : British Medical Journal, 2015. **350**: p. h391.
8. Drummond, M., Claxton, M., O'Brien, K., Stoddart, G.L., Torrance, G.W., *Methods for the economic evaluation of health care programmes*. 2015, Oxford, United Kingdom: Oxford Medical Publications.
9. Brazier, J.R., Salomon, J.A., Tsuchiya, A., *Measuring and valuing health benefits for economic evaluation*. 2017, Oxford, United Kingdom: Oxford University Press.