Clinical and cost-effectiveness of an in-home personalised health promotion intervention enabling independence in older people with mild frailty (HomeHealth): A Randomised Controlled Trial

Written by Louise Marston, Marie Le Novere and Rachael Hunter with input from Rachael Frost (Trial Manager), Yolanda Barrado-Martin (process evaluation) and Kate Walters (CI)
Version 5.0 Last updated on 18/09/2023
Version 4.0 Updated on 24/07/2023
Version 3.0 Updated on 30/11/2022
Version 2.2 Updated on 29/11/2022
Version 2.1 Updated on 28/11/2022
Version 2.0 Updated on 04/10/2022
Version 1.4 Updated on 04/10/2022
Version 1.3 Updated on 27/09/2022
Version 1.2 Updated on 23/09/2022
Version 1.1 Updated on 07/02/2022
Version 1.0 Updated on 21/09/2021
Version 0.5 Updated on 20/09/2021
Version 0.4 Updated on 15/09/2021
Version 0.3 Updated on 06/09/2021
Version 0.2 Updated on 09/08/2021
Version 0.1 Updated on 14/06/2021

Introduction

This analysis plan sets out the methods of analysing the predetermined primary, secondary and health economic outcomes for HomeHealth, which will be reported in the National Institute for Health Research, Health Technology Assessment report at the end of the trial and also in the main peer review paper(s) to result from this randomised controlled trial (RCT).

The analysis and reporting of this RCT will conform to the CONSORT, CONSORT extension for non-pharmacological interventions[1-4] and CHEERS statements[5] and the appropriate standard operating procedures written by Priment Clinical Trials Unit.

Further information on this trial can be found in the protocol version 7.0 (07/11/2022). The protocol is stored on: S:\FPHS_Priment_CTU\Projects\Current\Non CTIMPS\HomeHealth\6. Protocol\HomeHealth_interventional_protocol_priment_V7.0 07.11.22.

Trial summary

Aim

To test the clinical effectiveness of HomeHealth versus treatment as usual (TAU) in maintaining independence in an individual patient RCT.

Objectives

To determine whether there is a difference in independent functioning (our primary outcome) between those randomised to HomeHealth versus those randomised to TAU.

To evaluate whether there is a difference between those who are randomised to HomeHealth and those randomised to TAU in our secondary outcomes:

1) Ability to complete instrumental activities of daily living
2) Fried frailty phenotype
3) Wellbeing
4) Psychological distress
5) Loneliness
6) Cognition

7) Falls
8) Mortality
9) Carer burden

Study population

Inclusion criteria

Participants will be included if they:

- Are older people aged 65+
- Are registered with a general practice within the participating site area
- Score as 5 (mild frailty) on the Rockwood Clinical Frailty Scale (CFS)
- Are community-dwelling (including extra care housing)
- Have a life expectancy of more than six months
- Have capacity to consent to participate

People with dementia will not be excluded from the study, providing they fit the above criteria.

Exclusion criteria

Participants will be excluded if they:

- Are care home residents
- Have moderate to severe frailty (6 to 9 on CFS) or not frail (1 to 4 on CFS)
- Are receiving palliative care
- Are already case managed
- Lack capacity to consent

Trial design

Single-blind two-arm individually randomised trial comparing the HomeHealth intervention to TAU in older people with mild frailty at three sites. Data will be collected at baseline (prior to randomisation), six and 12 months' post-randomisation by blinded outcome assessors, as well as from medical notes after the 12-month assessment. Participants will be consented to be approached for further research, including longer-term (post-trial) follow-up using routinely collected data from NHS and local authority social care.

Randomised treatments

Intervention

HomeHealth is an individualised multi-domain behaviour change intervention based upon evidence and theory, which has been co-designed with stakeholders. Participants are initially offered up to six individual one-to-one sessions with a support worker over six months, and, where needs are more complex, more sessions (up to maximum 12) can be offered within this period. The first session will be face-to-face where possible, in accordance with current Government COVID-19 guidelines and using any personal protective equipment or social distancing measures. Subsequent sessions will be delivered face-to-face, by videoconferencing or by telephone according to participants' needs and preferences. If it is not possible to deliver the intervention face-to-face, potential participants will be offered the opportunity to defer enrolment in the trial until a later date or carry out all sessions by videoconferencing or telephone.

Core domains covered by the intervention include mobility (physical activity, exercise and falls prevention), under-nutrition or risk of malnutrition, mood (depression/ anxiety) and social engagement, with the potential for participants to include additional goals (for example, modifying their home environment). In each session, participants set and address self-directed independence and wellbeing goals, supported by a HomeHealth support worker through education, skills-training, overcoming barriers, providing feedback, maximising motivation, coping with setbacks and promoting habit formation. The support worker undertakes an initial behavioural assessment, considering the participant's capability,

opportunity and motivation to change, and their overall outcomes goals are broken down into behavioural goals and SMART objectives (Specific, Measureable, Achievable, Realistic, Timely). This assessment can include strategies to compensate for common problems causing barriers to change in this population; for example, fatigue, urinary incontinence. The support is individually tailored, and for frailer individuals or those with cognitive impairment this may include involving another person (for example, family member or friend), or providing practical support to overcome barriers, such as technology or provision of aids. Baseline function (capability) is taken into account – for example the exercise/ physical activity programme (exercises and intensity) will be tailored to ability and falls risk. Subsequent sessions then include reviewing goals and progress, addressing problem solving, coping with setbacks and low motivation, modifying or developing new goals as needed, forming an action plan and maintaining behavioural changes.

The service is delivered by a trained support worker who has experience working with older people, but without specialist qualifications. They are based within community/ voluntary sector teams working with older people. Where appropriate, support workers will encourage or enable participants to access local services (for example, falls prevention schemes, psychological therapies, hearing/ low vision aids, continence services, transport, dieticians, memory clinics, debt/ housing/ benefits advice, etc).

Control
Treatment as usual.

Sample size
We have calculated the sample size using the Minimal Clinically Important Difference (MCID) for the Modified Barthel Index (BI) at 12 months (1.85)[6]. We anticipate average functioning to decline over time without intervention in those with mild frailty; in our feasibility study scores declined by >1 point in 6 months in the control arm, and improved in our intervention arm[7]. If this decline is prevented, we would therefore expect a larger difference at 12 months than observed at 6 months in our feasibility study. The standard deviation (SD) was 3 for the BI in our feasibility RCT. This has been reported in other studies[8], but larger SDs have been reported in other settings in frail populations[9, 10]. We have therefore conservatively assumed an SD of 5 for our full trial, which would require 308 people (154 per group), with 90% power and 5% significance level. Whilst attrition was minimal (6%) at 6 months in our feasibility study, other studies have had higher attrition rates with longer follow-up[11].

We anticipate that clustering by therapist will be minimal and non-significant. No trials in those with mild frailty have reported therapist clustering (intraclass correlation coefficients (ICCs)), only clustering by GP practice in cluster RCTs in older community-based general populations[12]. Unpublished data from a PhD studying therapist effects in a secondary analysis of a cluster exercise trial in older people[13] suggested no significant clustering by therapist (ICC 0.01, P=0.54), so we have not inflated for therapist clustering. We anticipate employing seven HomeHealth Workers (HHWs) to deliver the intervention across three sites, with cluster sizes of 20-35 as HHWs are part-time.

Based on these estimates, a sample size of 386 people (193 per arm) is required to provide 90% power at the 5% significance level (two-sided) to detect an MCID of 1.85-points in the BI, assuming a 20% attrition rate at 12 months. The sample size calculation was performed using Stata.

Randomisation
Block randomisation, stratified by site, will be performed using a remote computerised system with allocation 1:1, described in a separate randomisation protocol. Random allocations will be issued on completion of the baseline eligibility criteria. There will be no replacements for participants who drop out or otherwise cannot comply with study procedures.

Blinding
Outcome assessors will be blind to allocation. The trial statistician performing the final analysis and the health economists will analyse the data blind to allocation and will be unblinded when results are confirmed. The statistician working with the Data Monitoring Committee may be unblinded during the course of that work. If that occurs, they will no longer attend Trial Management Group meetings.

Outcomes
The main analyses will be on data collected at 12 months.

Primary outcome
Modified Barthel Index (BI) at 12 months[14]. The BI is an interviewer-administered continuous scale from 0 to 100, where 100 reflects independent functioning in basic activities of daily living. It consists of 10 items with five levels of functioning from unable to perform the task to fully independent. Unable to perform the task always scores 0. Fully independent scores between 5 and 15 depending on the item. Likewise, the intermediate levels of functioning differ in scoring depending on the item. As this measure is interviewer-administered we anticipate that missing data will be very low.

Secondary outcomes
Nottingham Extended Activities of Daily Living (NEADL)[15-17]. It consists of 22 items in four sections; mobility, kitchen, domestic and leisure. For each item there are four responses; able, able with difficulty, able with help and unable/ did not do, scored 3, 2, 1, 0 respectively. Item scores are summed to give a possible score between 0 and 66. Those with missing items will not be able to have a total score calculated.

Items on the NEADL are scored in reference to activities performed in the 'last few weeks'. It is possible that some items will not have been carried out in the required time frame due to COVID-19 restrictions (eg. go out socially). The CRF has been updated to be able to distinguish between not doing an item due to restrictions and not doing it for another reason.

Fried Frailty Phenotype score[18] to assess for progression of frailty, including the following five components:

1. Gait speed - self-reported according to Op het Vald's (2018) questionnaire[19], consisting of 4 questions. The question "Do you walk more slowly than you'd like?" is given a score of 2 and the other three questions are given a score of 1 each if the response is yes. All questions are given a score of 0 if the response is no, giving an overall score of 0 to 5. Participants are considered frail if they score 3 to 5 and not frail otherwise.

2. Grip strength - self-reported according to Op het Vald's (2018) questionnaire[19], consisting of 2 questions. The question "Do you have trouble watering plants with a spray bottle?" is given a score of 2 if the response is yes, and "Do you feel like you have less hand strength than other people your age?" is given a score of 1 if the response is yes. Both questions are given a score of 0 if the response is no, giving an overall score of 0 to 3. A score of 0 is considered not frail and a score of 1 to 3 is frail. If the response to the second question is missing, this will be conservatively considered to be 0.

3. Physical activity – using the International Physical Activity Questionnaire-Elderly (IPAQ-E)[20], quantified according to the IPAQ-E guidelines[21]. It is structured so that it gives separate scores in three domains: walking, moderate intensity activity and vigorous intensity activity. It is scored by multiplying the number of minutes the domain is carried out by the frequency per week. If the number of minutes is less than 10, this is not included (minutes are set to 0). Number of minutes are further multiplied by the MET (metabolic

equivalent of task) minutes to get the score in terms of MET minutes. There are standard MET minutes for each domain[22]: walking = 3.3 MET minutes, Moderate physical activity = 4.0 MET minutes and Vigorous physical activity = 8.0 MET minutes. From these a total score in terms of MET minutes will be calculated. Participants are considered to be frail if participants are inactive or are not active enough to meet the criteria for minimally active, which is failure to meet any of these three criteria:

- 3 or more days of vigorous activity of at least 20 minutes per day **OR**
- 5 or more days of moderate-intensity activity or walking of at least 30 minutes per day **OR**
- 5 or more days of any combination of walking, moderate-intensity or vigorous intensity activities achieving a minimum of at least 600 MET-min/week.

4. Exhaustion – using the two exhaustion questions from the 7-item Centre for Epidemiological Studies Depression Scale: "7. I felt that everything I did was an effort" and "20. I could not get going"[18]. These are each scored 0 to 3. Participants are considered frail if they score 2 or 3 for either statement and not frail otherwise.

5. Weight loss using Weight loss during last three months from the Mini-Nutritional Assessment (MNA) Short Form[23]. This is scored:
   0 = Weight loss greater than 3 kg (6.6 pounds)
   1 = Does not know
   2 = Weight loss between 1 and 3 kg (2.2 and 6.6 pounds)
   3 = No weight loss
   Participants are considered to be frail if they have weight loss greater than 3kg (score 0).

To construct the Fried Frailty Phenotype score
The results of each of the five domains are frail (1) or not frail (0). These are summed, giving a total possible score of 0 to 5. Within that, 0 is considered not frail, 1 to 2 pre-frail and 3 to 5 as frail. To minimise the number of misclassifications, only one missing value is allowed when a person had a valid Fried score of 0–2. If a person had a valid Fried score of 3 points or more, two missing values are allowed, because this would not cause misclassification.

Gait speed, grip strength and physical activity using the IPAQ-E are separate outcomes in addition to being used as part of the Fried Frailty Phenotype score. IPAQ-E will be summarised as MET minutes per week and amount of sedentary time per week. All participants will have data from self-report for gait speed and grip strength; these give a score between 0 and 5 for gait speed and 0 and 3 for grip strength. These will also be categorised to frail versus not frail as described above. Where it has been possible to visit participants face-to-face they will also have physical grip strength (by dynamometer, kg) and gait speed (m/s) measures.

Wellbeing – using the Warwick-Edinburgh Mental Wellbeing Scale (WEMWBS)[24]. This is a 14 item scale, with five responses 1 (none of the time) to 5 (all of the time). Scores for each item are summed to give a score between 14 and 70 with a higher score indicating better mental wellbeing. If there are fewer than three missing items, within participant mean imputation of missing items will be carried out. If there are three or more missing items, the WEMWBS will not be scored.

Psychological distress – using the General Health Questionnaire-12 (GHQ-12)[25]. This consists of 12 items, each with four possible responses which are scored 0, 1, 2 or 3 (Likert method) with 0 indicating no psychological distress and 3 indicating the most psychological distress (worse than usual). These are summed to give a possible score between 0 and 36. If there are missing items, these will be scored as low score (0)[26].

Loneliness - using the University of California, Los Angeles 3-item loneliness scale (UCLA-3)[27]. It consists of three items, with three responses; hardly ever, some of the time, often which are scored 1, 2 and 3 respectively. The score of the items is summed to give a possible score between 3 and 9, with higher scores indicating greater loneliness. If there are missing items, the measure will not be scored.

Cognition – using the Montreal Cognitive Assessment (MoCA)[28] or telephone MoCA (T-MoCA)[29, 30]. The MoCA has a number of tasks, both visual and non-visual which attract differing scores. The measure is scored between 0 and 30 with 30 indicating no cognitive impairment. The T-MoCA consists of the items from the MoCA which do not require pencil and paper or other visual stimuli. The maximum score on the T-MoCA is 22. If any items are missing, the MOCA will not be scored.

The MoCA will be prioritised over the T-MoCA. However, if many assessments are conducted remotely, then we will use data from both measures to analyse the T-MoCA as an outcome (selecting the pertinent questions from the MoCA for this).

Falls (using the ProFANE consensus definition[31]). This will be number of falls in the previous six months, number of falls where an ambulance was called in the last six months and number of falls where the participant went to hospital in the last six months.

Mortality will be yes/ no by 6 and 12 months. It will be collected via proxy report (researchers attempting to contact the participant or via medical notes extraction).

Carer burden[32] using the iMTA Valuation of Informal Care Questionnaire (iVICQ). This asks questions in relation to whether the participant is a carer. The items are not aggregated, so will determine:

- n/N (%) who are carers
- mean (SD) or median (IQR) as appropriate number of hours caring each day
- carer burden using a numerical rating scale between 0 and 10
- How much longer the participant could continue giving care (categorical: <1 week, >1 week and <1 month, >1 month and <6 months, >6 months and <1 year, >1 year and <2 years, >2 years).

Data collection
Data collected at each time point are summarised in Table 1.

Table 1: Data collected at each time point

	Screening (phone)	**Screening (baseline)**	**Baseline assessment [1]**	**6 months**	**12 months**	**Medical notes extraction**
Visit No:		1	1	2	3	n/a
Window of flexibility for timing of visits:				-2 to +4 weeks	-2 to +4 weeks	-6 to 12 month period
Eligibility confirmation	X	X				
Informed Consent			X			
Demographics			X			
Alcohol (AUDIT-C)			X			
Smoking			X			
Index of Multiple Deprivation based on postcode			X			
Covid-19 status			X	X	X	
Modified Barthel Index (BI)			X	X	X	
Nottingham Extended Activities of Daily Living (NEADL)			X	X	X	
Gait speed			X	X	X	
Grip strength			X	X	X	
Physical activity (IPAQ-E)			X	X	X	
Weight loss			X	X	X	
Exhaustion			X	X	X	
Warwick-Edinburgh Mental Wellbeing Scale (WEMWBS)			X	X	X	
Euro-Qol 5D-5L (EQ-5D-5L)			X	X	X	
ICECAP-O			X	X	X	
12-item General Health Questionnaire (GHQ-12)			X	X	X	
UCLA 3-item			X	X	X	
Montreal Cognitive Assessment (MoCA)			X	X	X	
Falls (ProFANE consensus criteria)			X	X	X	
Mortality				X	X	X
Carer burden			X	X	X	
Client Services Receipt Inventory			X	X	X	
Randomisation			X			
Researcher perception				X	X	
Healthcare resource use						X
Comorbidities						X
Concomitant Medication review						X
Adverse Events review				X	X	X

Data entry

Data will be entered using a web based system set up by Sealed Envelope[33]. This has been set up so that it mirrors the data collection sheets in order. It also has range checks, consistency checks and for closed questions gives a number of options plus "other" where appropriate. Researchers who will be entering the data will have no access to the group allocation through this system. Data will be checked by a statistician and health economist before analysis and any problems reported to the Trial Manager, who will rectify them as appropriate before database locking and data analysis.

Statistical analyses

Estimands

Table 2 shows the estimand parameters as described in ICH E9 R1[34] with application to HomeHealth.

Table 2: Estimands considered in the analysis of HomeHealth

Estimand	Description in relation to HomeHealth data
Population	Those recruited to HomeHealth
Treatment condition	HomeHealth intervention from HomeHealth worker
Outcome	Modified Barthel Index at 12 months
Intercurrent event	Death before outcome – while alive Missing data due to non-completion – treatment policy Additional medications/ treatments – treatment policy
Population Level summary measure	Difference in modified Barthel Index means between randomised groups

Interim analyses

There are no planned analyses. However, this does not preclude the DMC from requesting interim analyses.

Final analyses

The primary analyses will be complete case, intention to treat (defined as all patients randomised, analysed according to their randomised group regardless of treatment received).

The CONSORT[1-4] flow diagram will be constructed by/ in collaboration with the Trial Manager who will have logs of patients who do and do not agree to take part in the study. It will include number of patients randomised to each arm of the trial, and the numbers who have follow up data available in each group.

Data will be analysed and checked using Stata v17[35] or higher.

Descriptive statistics

Initial analyses will describe summary statistics for all variables, both overall and by randomised group. Summary statistics for continuous variables will be mean, median, SD, lower quartile, upper quartile and reported appropriately according to distribution. Categorical variables will be described using frequencies and percentages. We will also summarise the distribution of intervention sessions attended and the baseline characteristics of those followed up and those not.

Analysis of the primary outcome

The primary outcome will be analysed using linear mixed modelling including the measures at 6 and 12 months in the outcome. The model will include Barthel Index score at baseline, site (the stratification variable, 3 sites), time, randomised group and an interaction between time and randomised group as fixed effects and participant as a random effect. Assumptions will be checked and if these are violated (eg non-Normal residuals), appropriate transformation or analogous modelling will be carried out. Results will be presented as coefficient (95% CI) and p-value.

The intraclass correlation coefficient will be reported, using HomeHealth worker as the unit of clustering in the intervention group.

Analysis of the secondary outcomes

Continuous secondary outcomes at 6 and 12 months (NEADL, Fried Frailty Phenotype score, continuous components of the Fried (gait speed, grip strength and IPAQ-E), WEMWBS, GHQ-12, UCLA-3, MoCA, T-MoCA) will be analysed in the same way as the primary outcome. Coefficients (95% CI) for BI at 6 months will be extracted from the primary analysis model.

Dichotomous outcomes (death and exhaustion) will be analysed using logistic regression.

Falls will be analysed using Poisson regression or an analogous alternative if the assumptions for Poisson are not met. If it is not possible to use Poisson regression or an analogous alternative, then we will consider dichotomising falls (into at least one fall versus no falls) and analysing in a similar way to other dichotomous outcomes. If there are too few events to perform statistical modelling for falls or the other dichotomous outcomes, they will be reported descriptively.

(Serious) adverse events, carer burden, hours caring and length of time the carer can continue caring for (from the iVICQ) will be reported descriptively.

Missing data

Baseline predictors of missingness of the primary outcome will be examined. If any are statistically significant, they will be included in a sensitivity analysis to restore the assumption of missing at random. Multiple imputation will not be carried out.

Sensitivity analyses

A number of sensitivity analyses will be carried out for the primary outcome (Modified Barthel Index at 12 months).

- Including predictors of missingness in an analogous model to the primary model.

- Carrying out mixed modelling, including HomeHealth worker (N=7) as a random effect to see if there is an impact of clustering through HomeHealth worker. Those in the treatment as usual group will be assigned codes so that they equate to clusters of one participant. Fixed effects in the model will be the same as the primary outcome model.

- We will perform a complier average causal effects (CACE)[36] analysis after unblinding of the statisticians using a threshold dosage of 3+ sessions for compliance to determine the average treatment effect of participants who would have adhered to the protocol regardless of how they were randomised. We will use ivregress in Stata.

- We will perform the delta based imputation method[37] if there is a difference of more than 5% missing data between study arms at 6 months and 12 months. We will exclude those who die before 6 months to maintain the while alive intercurrent event in the estimand framework (Table 2). The imputation model will include the variables in the substantive model, baseline variables associated with missingness of the outcome and baseline variables associated with the outcome. If the following are not associated with the outcome or missingness, they will also be included based on clinical knowledge of this participant group: number of long term conditions (Cambridge multi-morbidity score), number of medications (polypharmacy), MoCA score, Fried Frailty Phenotype score, the participant's carer status, baseline hospitalisations. The substantive model will be similar to the primary outcome in form, using Rubin's rules[38]. The delta will be 0 for the intervention group and -3 for the treatment as usual group.

Sensitivity analysis for the NEADL
Analysing the data excluding participants who indicated that their performance in any item was affected by COVID-19.

Process evaluation
The process evaluation analyses are indicative. They are hypothesis generating only as they are not powered to give definitive results.

1. To explore the reach of the intervention we will descriptively compare the baseline socio-demographic characteristics of participants (age, sex, ethnic group, country of birth (UK versus another country), education, IMD, housing tenure) against the data from the electoral ward the participants were recruited from in the UK Census 2021 (where available, if not then from 2011) or the IMD available from Public Health England.

2. In addition to the CACE analysis, we will determine whether those who get the therapeutic dose of the intervention (3+ sessions) have higher BI at 12 months than those who do not. Number of sessions will be dichotomised to 3+ versus less than three. Those in the TAU group will be coded as having less than three sessions. There will be an interaction between number of sessions indicator and randomised group. This will be analysed in a similar way to the primary analysis including baseline BI score.

3. To assess whether choice of outcome goal type (mobility, nutrition, psychological, social, cognitive, other) as reported in Health and Wellbeing plans is associated with BI scores at 12 months, similar analyses to 2 will be carried out with an interaction between goal mobility (yes/ no), nutrition (yes/ no), psychological (yes/no), social (yes/no), cognitive (yes/no), other (yes/ no) and randomised group. Each intervention participant can work on more than one outcome goal over time, which would be reflected through more than one Health and Wellbeing Plan being completed over sessions containing different information in the outcome goal section. If this is the case goal types will be aggregated (eg, if first the participant decides to work on mobility and later on a nutrition type of goal, both mobility and nutrition will be marked). Even when a participant decides to work on an outcome goal throughout the whole of the intervention, the goal may reflect different types of goals (e.g., be able to walk further to go and visit my friend – which would be coded as mobility and social), so the outcome goal types are not mutually exclusive, combinations of goal types can occur. In these cases, participants will be included in multiple analyses. Modelling will be undertaken separately for each goal if there are sufficient participants selecting a given type of goal. Those in the TAU group will be coded as no. If there are sufficient numbers selecting a given goal type, we will explore whether there are effects on the most related secondary outcome to assess whether specific behavioural targets show potential for effectiveness: mobility (gait speed and IPAQ-E score), nutrition (weight loss category from the MNA questionnaire; likely to be dichotomised to weight loss of at least 3kg versus less/ not known), psychological (GHQ-12), social (UCLA-3), cognitive (MoCA/ t-MoCA).

4. As goal setting is a key component of HomeHealth, we will explore whether overall progress to meeting SMART goals over appointments (reflected and rated by the HomeHealth worker at the bottom of the HomeHealth workers' fidelity checklists) using goal progress rated on a score of 0-2 (where 0= no progress, 1= partial progress and 2 = goal met), to see whether there is greater impact on BI. In the intervention group, if a SMART goal is rated across several appointments, we will explore the last rating given for each SMART goal. Where a SMART goal is introduced but not rated at any point (e.g., is added just before withdrawing, there is no further follow-up on that goal, etc), it will be excluded from the analysis as for those participants who decide against setting any goals. As it is expected that more than one SMART goal will be set by most participants over the course of the intervention, but there is no minimum or maximum amount permitted, we will

average progress towards several SMART goals so that each participant has one average goal progress measurement. Goal progress will be kept as a continuous variable. Goal attainment scaling (progress towards the overall outcome goal, with scale parameters agreed when the outcome goal was set and final progress ratings made at the final appointment) will be used if sufficient data are available. In this case, we will combine scores, without weighting, from all outcome goals to give a score. In a similar way to `before, we will model the BI using linear regression with an interaction between progress towards goals measured by SMART goals or GAS and randomised group. In the TAU group, the progress towards goals will be set at 0 as there were no goals set, so no progress was made.

5. The number of appointments completed, their duration and the appointment format (e.g., face to face, online or telephone) will be logged and summarised with appropriate statistics.

6. Intervention fidelity will be calculated as an overall score derived from the researcher-assessed fidelity checklists applied to appointment transcripts (from 19 randomly selected participants, with a minimum of 3 appointments audio-recordings available per participant), which scores 10-17 items (depending on appointment) as completed, completed to some extent, not done, not applicable and unable to assess. The proportion of items completed or completed to some extent out of those that are applicable (that is completed, completed to some extent and not done) will be calculated and summarised overall and by appointment type (first/ subsequent/ final), site and HomeHealth worker. HomeHealth workers also complete self-reported versions of the same fidelity checklist. We will summarise fidelity across all appointments in the same way as with researcher-assessed fidelity checks and calculate agreement between those which are self-reported (by HomeHealth workers) and independently (by two researchers) rated. The level of agreement will be calculated.

7. If sufficient remote delivery takes place (most likely to be used as COVID-19 mitigation), the impact of remote delivery will be explored. Here, remote delivery means all sessions delivered remotely. We will record how many sessions are being delivered face-to-face and remotely.

Health economic analysis

Aim

The primary aim is to calculate the mean incremental cost per quality adjusted life year (QALY) using the EQ-5D-5L and the relevant UK tariff of the HomeHealth intervention compared with treatment as usual (TAU) from a health and social care cost-perspective.

Secondary aims:

Calculate the mean incremental cost per years of full capability (YFC) gained using the ICECAP-O and its respective tariff for the durations of the trial from a health care cost perspective.

Additional secondary analyses will report the mean incremental cost per QALY and YFC gained from a wider cost perspective.

Outcomes

EQ-5D-5L[40, 41] The EQ-5D-5L is a measure of health related quality of life with an associated preference based tariff that can be used to calculate QALYs. The EQ-5D-5L will be collected at baseline, 6 and 12 months' follow-up. Utility weights will be calculated from responses to

the EQ-5D-5L using the van Hout EQ-5D-3L mapping algorithm[42]. Utility weights calculated from the UK tariff published by Devlin et al[43]will be included in a sensitivity analysis.

ICE pop CAPability measure for Older people (ICECAP-O)[44] Information on capability will be collected using the ICECAP-O, a measure focused on wellbeing rather than solely health, specifically designed for older people. Years of full capability will be calculated using the UK index values developed by Coast et al[45] and responses to the ICECAP-O.

Health care resource use - Electronic Medical Records Data on primary and secondary healthcare usage, medications will be collected by a researcher from primary care electronic medical records (EMRs) covering the last 18 months (at least 12 months' post baseline and up to 6 months pre baseline).

Client Service Receipt Inventory (CSRI) Resource use not covered by primary care medical records will be collected using modified version of the CSRI previously tested in the feasibility trial. The CSRI collects information on NHS and private Healthcare contacts over the previous 6 months, over the counter medicines, ambulance and hospital use following falls, residential and nursing care, personal care and help at home, services for transport, benefits, and caring responsibilities. This is collected at baseline, 6 and 12 months asking about the previous 6 months.

Cost data

Cost of HomeHealth

The cost of the HomeHealth intervention will be taken by multiplying the unit cost of the support worker delivering the intervention by the duration and number of sessions attended by each participant. This will be variable as participants were offered six sessions, but can have up to 12 sessions depending on the complexity of their needs. The total cost of training and supervision of the support workers will be divided by the number of participants in the treatment group. We will report the mean cost and standard deviation in the treatment group.

Resource use

Descriptive statistics for the percentage of participants using each item in the CSRI and the volume of usage will be reported at baseline, 6 and 12 months by group. Information on data completeness will also be reported.

Cost of health and social care resource use

The cost of health and social care resource use for the HomeHealth group versus TAU will be calculated using the data collected in the modified CSRI and from primary care EMRs. These will be calculated for each participant using published unit costs from the most recent version of the Unit Costs of Health and Social Care by the Personal Social Services Research Unit (PSSRU), NHS reference costs and other published sources.

Mean cost per participant will be reported by group as total cost (6 months plus 12 months) per participant and by service use type. The difference in health and social care costs between the groups will be reported. Mean incremental cost will be calculated using regression analysis adjusting for baseline cost and site. Bootstrapping will be used to calculate 95% CIs.

Wider Societal Costs

Wider societal costs will include the cost of private services, over the counter medicines and vitamins, voluntary sector, services for local transport and carer time divided into state funded, privately funded and unpaid carers (family, friends and close others). Cost of carer time will be costed as the hourly cost to provide social care. It is generally recommended that government funded welfare payments, sometimes called "transfer costs" are excluded from economic evaluations as they are cost neutral to society, but that there are some instances where decision makers might be interested in their inclusion[46]. As a result, government funded

welfare payments will be included in a sensitivity analysis. For those who have reported they receive payments but have not disclosed the amount we will take average values from the gov.uk website.

Quality-adjusted Life Years (QALYs)
QALYs will be calculated using the area under the curve method adjusting for baseline utility and site. Mean utility values and mean unadjusted QALYs from baseline to 12 months will be reported for both groups. The incremental mean difference in QALYs will be calculated using regression analysis adjusting for baseline, site. This will be reported for both groups with bootstrapped 95% CIs.

Capability and Years of full Capability (YFC) Baseline Capability score and unadjusted YFC at 12 months will be reported for both groups. The incremental mean difference in YFC between the HomeHealth group and TAU will be calculated using regression analysis adjusting for baseline capability score and site.

Discounting
No discounting will be included as the time horizon is 12 months.

Primary Analysis
Incremental cost-effectiveness Ratio (ICER) We will report mean incremental cost per QALY and YFC gained between the HomeHealth intervention and TAU at 12 months. Costs will be bootstrap adjusted as specified and will include the cost of health and social care resource use in the intervention and TAU arm and the cost of HomeHealth in the intervention arm only. Seemingly unrelated regression will be used to account for the correlation between costs and QALYs.

Cost-effectiveness acceptability curve (CEAC) and cost-effectiveness plane (CEP) The bootstrapped means and 95% CIs for costs, QALYs and YFC will be used to calculate the probability that the HomeHealth intervention is cost-effective compared with TAU for a range of cost-effectiveness thresholds for one QALY gained and one additional YFC. A cost-effectiveness plane and CEAC will be reported for the bootstrapped results.

Missing data
In line with the statistical analysis plan, we will investigate predictors of missingness. If any are statistically significant, we will include them in a sensitivity analysis to restore the assumption of missing at random. If there is greater than 10% of participants missing from the analysis due to missing data, we will consider multiple imputation for health economic outcomes as recommended by Faria et al[47].

Sensitivity analyses
We will conduct sensitivity analysis testing the impact of changing assumptions around the cost of HomeHealth such as:

- The number of participants per HomeHealth support worker.
- The frequency of supervision.
- The number of recommended sessions per patient.

The following sensitivity analyses will also be conducted:
- As there may be more loss to follow-up for the CSRI, including costs collected from primary care EMR only.
- Including government funded welfare payments alongside wider societal costs.
- A range of sensitivity analyses to evaluate the impact of missing data, using progressively more complicated models.

Secondary Analyses

Cost-effectiveness from a societal perspective

We will report the ICER, CEAC and CEP for both QALYs and YFC for HomeHealth versus TAU at 12 months from a societal perspective. Societal costs will include private healthcare, pocket costs including out of pocket costs, voluntary sector services and the cost of paid and unpaid carers. QALYs will be calculated using the Value Set for England as well as the EQ-5D-3L mapping algorithm.

In addition to the main analysis described in this plan we will undertake a budget impact analysis including long-term analyses/decision models. The analysis plans for this component of the project will be detailed in a separate document.

References

1 Begg C, Cho M, Eastwood S, Horton R, Moher D, Olkin I, et al. Improving the quality of reporting of randomised controlled trials. The CONSORT statement. JAMA 1996;276:637-9.

2 Moher D, Schulz KF, Altman DG, CONSORT Group. The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomised trials. Lancet 2001;357:1191-4.

3 Schulz KF, Altman DG, Moher D. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. BMJ 2010;340:c332

4 Boutron I, Altman DG, Moher D, Schulz KF, Ravaud P, DJ C, et al. CONSORT Statement for Randomized Trials of Nonpharmacologic Treatments: A 2017 Update and a CONSORT Extension for Nonpharmacologic Trial Abstracts. Annals of Internal Medicine. American College of Physicians; 2017 Jul 4;167(1):40

5 Husereau D, Drummond M, Petrou S, Carswell C, Moher D, Greenberg D, Augustovski F, Briggs AH, Mauskopf J, Loder E Consolidated Health Economic Evaluation Reporting Standards (CHEERS) statement BMJ. 2013;346:f1049

6 Hsieh Y-W, Wang C-H, Wu S-C, Chen P-C, Sheu C-F, Hsieh C-L. Establishing the minimal clinically important difference of the Barthel Index in stroke patients. Neurorehabilitation and neural repair. 2007;21(3):233-8.

7 Walters K, Frost R, Kharicha K, Avgerinou C, Gardner B, Ricciardi F, et al. Home-based health promotion for older people with mild frailty: the HomeHealth intervention development and feasibility RCT. Health technology assessment (Winchester, England). 2017;21(73):1

8 Serra-Prat M, Sist X, Domenich R, Jurado L, Saiz A, Roces A, et al. Effectiveness of an intervention to prevent frailty in pre-frail community-dwelling older people consulting in primary care: a randomised controlled trial. Age and ageing. 2017;46(3):401-7.

9 Huguet LG, González MN, Kostov B, Carmona MO, Francia CC, Nieto MC, et al. Pre Frail 80: Multifactorial Intervention to Prevent Progression of Pre-Frailty to Frailty in The Elderly. The journal of nutrition, health & aging. 2018:1-9.

10 Romera-Liebana L, Orfila F, Segura JM, Real J, Fabra ML, Möller M, et al. " Effects of a primary-care based multifactorial intervention on physical and cognitive function in frail,

elderly individuals: a randomized controlled trial". The Journals of Gerontology: Series A. 2018:glx259

11 Puts MTE, Toubasi S, Andrew MK, Ashe MC, Ploeg J, Atkinson E, et al. Interventions to prevent or reduce the level of frailty in community-dwelling older adults: a scoping review of the literature and international policies. Age and ageing. 2017;46(3):383-92.

12 Smeeth L, Ng ES-W. Intraclass correlation coefficients for cluster randomized trials in primary care: data from the MRC Trial of the Assessment and Management of Older People in the Community. Controlled clinical trials. 2002;23(4):409-21.

13 Gawler S. An exploration of therapist and dose effects within an exercise intervention study for reducing falls and falls risk factors in community dwelling older people. . Doctoral Thesis, University College London. 2016.

14 Shah S, Vanclay F, Cooper B. Improving the sensitivity of the Barthel Index for stroke rehabilitation. Journal of clinical epidemiology. 1989;42(8):703-9.

15 Wu C-Y, Chuang L-L, Lin K-C, Horng Y-S. Responsiveness and validity of two outcome measures of instrumental activities of daily living in stroke survivors receiving rehabilitative therapies. Clinical rehabilitation. 2011;25(2):175-83.

16 Harwood R, Ebrahim S. The validity, reliability and responsiveness of the Nottingham Extended Activities of Daily Living scale in patients undergoing total hip replacement. Disability and rehabilitation. 2002;24(7):371-7.

17 Lin S-Y, Kerse N, McLean C, Moyes S. Validation of quality of life and functional measures for older people for telephone administration. Journal of Primary Health Care. 2010;2(1).

18 Fried LP, Tangen CM, Walston J, Newman AB, Hirsch C, Gottdiener J, et al. Frailty in older adults: evidence for a phenotype. The Journals of Gerontology Series A: Biological Sciences and Medical Sciences. 2001;56(3):M146-M57.

19 Op Het Veld LPM, de Vet HCW, van Rossum E, Kempen G, van Kuijk SMJ, Beurskens A. Substitution of Fried's performance-based physical frailty criteria with self-report questions. Arch Gerontol Geriatr. 2018;75:91-5.

20 Hurtig-Wennlof A, Hagstromer M, Olsson LA. The International Physical Activity Questionnaire modified for the elderly: aspects of validity and feasibility. Public Health Nutr. 2010;13(11):1847-54.

21 IPAQ. Guidelines for Data Processing and Analysis of the International Physical Activity Questionnaire (IPAQ) – Short and Long Forms. November 2005.

22 Ainsworth BE, Haskell WL, Whitt MC, Irwin ML, Swartz AM, Strath SJ, O'Brien WL, Bassett DR Jr, Schmitz KH, Emplaincourt PO, Jacobs DR Jr, Leon AS Compendium of Physical Activities: An update of activity codes and MET intensities *Medicine and Science in Sports and Exercise*, 2000;32 (Suppl):S498-S51

23 Rubenstein LZ, Harker JO, Salva A, Guigoz Y, Vellas B. Screening for undernutrition in geriatric practice: developing the short-form mini-nutritional assessment (MNA-SF). J Gerontol A Biol Sci Med Sci. 2001;56(6):M366-72.

24 Tennant R, Hiller L, Fishwick R, Platt S, Joseph S, Weich S, et al. The Warwick-Edinburgh Mental Well-being Scale (WEMWBS): development and UK validation. Health Qual Life Outcomes. 2007;5:63.

25 Goldberg DP, Gater R, Sartorius N, Ustun TB, Piccinelli M, Gureje O, et al. The validity of two versions of the GHQ in the WHO study of mental illness in general health care. Psychol Med. 1997;27(1):191-7.

26 https://support.gl-assessment.co.uk/knowledge-base/assessments/general-health-questionnaire-support/about-the-general-health-questionnaire/faqs/ accessed September 2022

27 Hughes ME, Waite LJ, Hawkley LC, Cacioppo JT. A Short Scale for Measuring Loneliness in Large Surveys: Results From Two Population-Based Studies. Res Aging. 2004;26(6):655-72.

28 Nasreddine ZS, Phillips NA, Bedirian V, Charbonneau S, Whitehead V, Collin I, et al. The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. J Am Geriatr Soc. 2005;53(4):695-9.

29 Zietemann V, Kopczak A, Muller C, Wollenweber FA, Dichgans M. Validation of the Telephone Interview of Cognitive Status and Telephone Montreal Cognitive Assessment Against Detailed Cognitive Testing and Clinical Diagnosis of Mild Cognitive Impairment After Stroke. Stroke. 2017;48(11):2952-7.

30 Pendlebury ST, Welch SJ, Cuthbertson FC, Mariz J, Mehta Z, Rothwell PM. Telephone assessment of cognition after transient ischemic attack and stroke: modified telephone interview of cognitive status and telephone Montreal Cognitive Assessment versus face-to-face Montreal Cognitive Assessment and neuropsychological battery. Stroke. 2013;44(1):227-9.

31 Lamb SE, Jorstad-Stein EC, Hauer K, Becker C, Prevention of Falls Network E, Outcomes Consensus G. Development of a common outcome data set for fall injury prevention trials: the Prevention of Falls Network Europe consensus. J Am Geriatr Soc. 2005;53(9):1618-22.

32 Hoefman R, Van Exel N, Brouwer W. iMTA valuation of informal care questionnaire (iVICQ). Retrieved. 2011;17(11):2013.

33 http://www.sealedenvelope.com/ accessed September 2021

34 International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use. Addendum on Estimands and Sensitivity Analysis in Clinical Trials to the Guideline on Statistical Principles for Clinical Trials. 2019 Available from https://www.ema.europa.eu/en/documents/scientific-guideline/draft-ich-e9-r1-addendum-estimands-sensitivity-analysis-clinical-trials-guideline-statistical_en.pdf

35 StataCorp. 2021. *Stata Statistical Software: Release 17*. College Station, TX: StataCorp LLC

36 Connell AM Employing Complier Average Causal Effect Analytic Methods to Examine Effects of Randomized Encouragement Trials Am J Drug Alcohol Abuse. 2009 ; 35(4): 253–259

37 Cro, S, Morris, TP, Kenward, MG, Carpenter, JR Sensitivity analysis for clinical trials with missing continuous outcome data using controlled multiple imputation: A practical guide. *Statistics in Medicine*. 2020;39:2815–2842.

38 Rubin DB. *Multiple Imputation for Nonreponse in Surveys*. Wiley Series in Probability and Mathematical Statistics. New York, NY: John Wiley & Sons; 1987

39 Turner-Stokes L Goal Attainment Scaling (GAS) in Rehabilitation A practical guide Goal Attainment Scaling in Rehabilitation (kcl.ac.uk) Accessed September 2021

40 Brooks R RR, De Charro F, editors. The measurement and valuation of health status using EQ-5D: a European perspective: evidence from the EuroQol BIOMED Research Programme.: Springer Science & Business Media; 2013.

41 Haywood KL GA, Fitzpatrick R. Quality of life in older people: a structured review of generic self-assessed health instruments. Quality of life Research. 2005;14(7):1651-68.

42 Van Hout B, Janssen M, Feng Y et al Interim scoring for the EQ 5D 5L: Mapping the EQ 5D 5L to EQ 5D 3L value sets. Value in Health, 2012 15: 708-15.

43 Devlin NJ, Shah KK, Feng Y, Mulhern B, van Hout B Valuing health-related quality of life: An EQ-5D-5L value set for England; Health Economics. 2018;27:7–22

44 Bulamu NB, Kaambwa B, Ratcliffe J. A systematic review of instruments for measuring outcomes in economic evaluation within aged care. Health Qual Life Outcomes. 2015;13:179.

45 Coast J, Flynn TN, Natarajan L, Sproston K, Lewis J, Louviere JJ, et al. Valuing the ICECAP capability index for older people. Soc Sci Med. 2008 Sep 1;67(5):874–82.

46 Frisman L. Rosenheck R. How transfer payments are treated in cost-effectiveness and cost-benefit analyses. Administration and Policy in Mental Health, 1996 23(6): p.533-46.

47 Faria R, Gomes M, Epstein D, White IR. A guide to handling missing data in cost-effectiveness analysis conducted within randomised controlled trials. Pharmacoeconomics. 2014;32(12):1157-70.