Statistical Analysis Plan

PROSPER III (PROductivity Study of Presbyopia Elimination among aRtisans: a mixed methods randomised trial on the effect of providing near glasses on workplace retention of Indian textile workers)

DOCUMENT HISTORY

Version	Date	Edited by	Comments	Timing in relation to unblinding
1.0	TBD	GM	Approved by signatories, updated to version 1.0	Prior to trial registration

CONTRIBUTORS

Nathan Congdon (QUB), Anant Nyshadam (UMICH), Ling Jin (ZOC), Lynne Lohfeld (QUB), Graeme MacKenzie (Clearly Initiatives)

1. TABLE OF CONTENTS

1. TABLE OF CONTENTS ..... **Error! Bookmark not defined.**
2. INTRODUCTION 3
3. PERSONNEL 3
Chief Investigator 3
Principal Investigators 3
Trial Statisticians 3
4. BACKGROUND INFORMATION 3
4.1 Rationale 3
4.2 Objectives of the trial 4
4.3 Trial design 5
4.4 Eligibility 5
4.4.1 Inclusion criteria 5
4.4.2 Exclusion criteria 5
4.5 Interventions 5
4.6 Definition of primary and secondary outcomes 5
4.6.1 Primary outcome 5
4.6.2 Secondary outcomes 6
4.7 Hypothesis framework 6
4.8 Sample size & power 6
4.9 Intervention allocation 6
4.10 Data collection schedule 7
4.10.1 Data collection before trial 7
4.10.2 Data collection during trial 7
4.11 Interim analyses and stopping rules 7
4.12 Trial reporting 8
5. PROTOCOL DEVIATIONS 8
5.1 Major 8
5.2 Minor 8
5.2.1 Participants randomised in error 8
5.2.2 Participants who do not receive allocated intervention 8
6. ADHERENCE TO THE INTERVENTION 8
7. ANALYSIS POPULATIONS 9
7.1 Primary analysis strategy 9
7.2 Descriptive analysis population 9
8. DESCRIPTIVE ANALYSES 9
8.1 Representativeness of trial population and participant throughput 9
CONSORT Flow Diagram 10
8.2 Baseline comparability of randomised groups 10
8.3 Losses to follow-up 11
8.4 Adherence to intervention 11
9. COMPARATIVE ANALYSES 11
9.1 Analysis of primary and secondary outcomes 11
9.2 Pre-specified subgroup analysis 12
9.3 Sensitivity analysis 12
9.4 Significance levels and adjustment of p-values for multiplicity 12
9.5 Procedure for accounting for missing, unused, and spurious data 13
9.6 Exclusion of data 13
9.7 Statistical software employed 13
10. QUALITATIVE SUB-STUDY 13
10.1 Research Design 13
10.2 Sample and Sample Size 13
10.3 Recruitment 14
10.4 Time Period 14
10.5 Data Collection 14
10.6 Data Analysis 14
10.7 Outcomes 16
10.8 Rigour 16
11. SAFETY DATA ANALYSIS 17
11.1 Serious adverse events 17
12. ADDITIONAL EXPLORATORY ANALYSIS 17
13. DEVIATION FORM ANALYSIS DESCRIBED IN PROTOCOL 17
13. REFERENCES 18
14. APPROVAL 19

2. INTRODUCTION

This document details the proposed presentation and analyses for the main analysis reporting results from the USAID and Clearly funded randomised controlled trial PROSPER III (PROductivity Study of Presbyopia Elimination among aRtisans: a randomised trial on the effect of providing near glasses on workplace retention of Indian textile workers). The results reported in this publication will follow the strategy set out here. Subsequent analyses of a more exploratory nature will not be bound by this strategy, although they are expected to follow the broad principles described. The principles are not intended to curtail exploratory analysis (for example, to decide cut- points for categorisation of continuous variables), nor to prohibit accepted practices (for example, data transformation prior to analysis); rather they are intended to establish the rules that will be followed, as closely as possible, when analysing and reporting on the trial and the qualitative sub-study.

The analysis plan will be available on request when the principal manuscripts are submitted for publication. Suggestions for subsequent analyses by journal editors or referees will be considered carefully and carried out, as far as possible, in line with the principles of this analysis plan.

Any deviations from the data analysis plan will be described and the rationale given in the final report of the trial. The analysis will be carried out by an identified, appropriately qualified and experienced statistician, who will ensure the integrity of the data during processing. Examples of such procedures include quality control and evaluation procedures.

3. PERSONNEL

Chief Investigator

Professor Nathan Congdon, Queen's University Belfast, UK

Principal Investigators

Professor Achyuta Adhvaryu, University of Michigan, USA

Professor Anant Nyshadam, University of Michigan, USA

Professor Nathan Congdon, Queen's University Belfast, UK

Trial Statisticians

Ling Jin, Zhongshan Ophthalmic Center, PRC

4. BACKGROUND INFORMATION

4.1 Rationale

PROSPER III is a mixed methods randomised controlled trial (RCT) assessing the effect of reading glasses on the retention of presbyopic textile workers in Bangalore, India. PROSPER

III's primary approach is a randomized controlled trial (RCT), and the secondary or embedded approach is a descriptive qualitative study.

4.2 Objectives of the trial

Globally, 3 billion people do not have the eyeglasses they need to earn, learn, travel safely in traffic and participate in civic life. Among these, 1.1 billion people lack a simple pair of reading glasses to correct impaired near vision, called presbyopia.[1] Presbyopia, the essentially universal decline in unaided near vision that occurs with aging, is the world's most common cause of vision impairment. Loss of accommodation (ability to change focus from distance to near) due to presbyopia can begin as early as age 30 years, commonly becomes functionally apparent by 40, and is essentially complete by 55, meaning that presbyopia is most common at the height of the working years. The correction of presbyopia is both safe and affordable: requiring nothing more than a pair of reading glasses to restore vision at near. Despite this, only approximately 10% of people in need of reading glasses in low and middle-income countries actually have them.

Data from the International Labour Organization shows that the rate of having a job among people over the age of 40 in low- and middle-income countries is declining. Our previous randomized trials in India[2] demonstrated that the provision of free reading spectacles significantly improved the productivity of agricultural workers on a tea plantation over the course of a harvest season. The PROSPER trial enrolled 751 adults (mean age 47 years) with uncorrected presbyopia in Assam, India, and showed significantly higher productivity among workers randomized to receive free glasses compared to Controls (21.7% relative productivity increase; effect size 1.01 [95% CI 0.86–1.16]; p<0.0001). Intervention-group compliance with study glasses reached 84.5% by closeout. Regression model predictors of greater productivity increase included intervention group membership (an extra 5.25 kg of tea leaves picked per day [95% CI 4.60–5.91], p<0.0001) and, among intervention participants, older age (p=0.039) and better compliance with the intervention (p<0.0001). PROSPER I revealed a significant interaction between age and study group for the main study outcome. Older participants in the intervention group had significantly greater productivity increases than younger participants. Older participants in the control group, having more pronounced, uncorrected presbyopia, were less able than their younger peers to take advantage of higher crop yields during the peak high season, resulting in lower productivity increases. This strong interaction of age and productivity with study group adds to the biological plausibility of the results of PROSPER I. To place these results into perspective, the relative productivity increase in the productivity of workers receiving reading glasses was as large as or larger than that reported for any other health intervention trial in low-income and middle-income countries.[3-5]

In light of the above findings, PROSPER III will assess whether free reading glasses are able to extend the productive working life of workers (i.e increase retention) in textile factories in Bangalore, India. We hypothesise that worker retention over the 18-month evaluation period will be greater in the Intervention compared to the Control group. Additional qualitative data will be collected to enhance understanding of factors affecting retention of study subjects (textile workers aged 30 years and above with uncorrected presbyopia who are employed by Shahi Exports Private Limited in Karnataka, India).

4.3 Trial design

PROSPER III is an investigator-masked, multi-center mixed methods randomized controlled trial.

4.4 Eligibility

4.4.1 Inclusion criteria

Shahi employees will be eligible to participate if:

- they are aged 30 years or older
- have an unaided distance visual acuity of 6/12 or better in both eyes
- have presbyopia, defined as the inability—correctable with reading glasses—to read the N8 line using both eyes together, on a tumbling E near vision chart at 40cm
- have worked in the sewing department for 3 months or more

4.4.2 Exclusion criteria

Shahi employees will be ineligible to participate if:

- they own reading or distance glasses (regardless of accuracy)
- have ocular pathology in either eye detected during the eye examination, or history of such disease based on self-report
- have a low likelihood of completing follow-up in the study due to current plans to move out of the area or leave employment at Shahi during the follow-up period

4.5 Interventions

Eligible participants will be randomly assigned to Intervention or Control Groups (1:1). Intervention group participants will receive free reading glasses within a week of undergoing a vision screening at the factories. In addition to receiving free reading glasses, participants in the Intervention Group will be eligible for free replacement glasses in the event of loss or damage throughout the trial. Control group participants will receive free reading glasses at the end of the assessment period (18 months after vision screening). An interim analysis will be performed by the Data Monitoring and Ethics Committee (DMEC). If the Trial Steering Committee (TSC) terminates the trial on the recommendation of the DMEC, then Control Group participants will receive free reading glasses at trial closure. The trial will be investigator-masked, but not participant-masked, because the investigators do not feel provision of zero-power glasses to the control group is ethical. However, participants will have limited knowledge of the study hypothesis, limiting potential placebo effects.

4.6 Definition of primary and secondary outcomes

4.6.1 Primary outcome

The primary objective of the study is to estimate the effect of the provision of free reading glasses on the retention of textile workers employed in Shahi's sewing department. The time-to-event endpoint for PROSPER III is defined as the time between randomization to

Intervention or Control Group and loss of employment at Shahi. Employment status for each participant will be assessed from trial entry to closure as recorded in Shahi's Human Resource Management Database.

4.6.2 Secondary outcomes

The secondary outcomes are:

- change in efficiency (efficiency is defined as the number of items completed divided by target number of items per unit time as defined by industry standards)
- proportion of worker leaving employment in the 30th percentile for mean efficiency for the two months of employment prior to leaving
- reasons given by supervisor and colleagues of former sew-ers (separately) for workers no longer being employed at Shahi
- proportions of sew-ers who state they are satisfied or very satisfied with work, feel valued or very valued at work, and are likely or very likely to stay with current employer
- change in self-assessed productivity
- self-report of having been approached by a glasses champion and/or a supervisor
- change in attitude towards eye health, glasses wear and uptake of eyecare services
- changes in skill grade and associated change in wage
- adherence with spectacle wear
- Quality of Life using the THRIVE Near Vision Quality of Life tool.
- intervention cost-effectiveness

4.7 Hypothesis framework

PROSPER III is a superiority trial comparing glasses wear to non-wear. Analysis of the trial will entail calculation of treatment effect measures and confidence intervals to assess the difference between the two arms.

4.8 Sample size & power

Previous retention trials in this setting have showed 18-20% of overall likelihood of attrition at 5 months follow-up, 50% overall likelihood of attrition at 12 months, 60% at 18 months, 70% at 24 months. Assuming a work retention rate of 50% at 1 year and 40% at 18 months in the Control group, we anticipate that our glasses intervention will increase retention by 20% in the Intervention group at 1 year (hazard of 0.8). Data analysis using a Cox proportional hazard model with 2-sided significance at the 5% level and 80% power yields a total sample size of 1,260 (630 in each group). No correction for attrition is required as attrition is the outcome and therefore no attrition from the sample is possible.

4.9 Intervention allocation

Consenting participants eligible for the trial will be divided into six strata according to age (<median, ≥median), work tenure at the textile factory (<median, ≥median), and efficiency during baseline assessment (<median, ≥median). Participants in each stratum will be randomized 1:1 with block size of four to either the Intervention or Control Groups. The randomisation sequence will be generated by the study statistician at the Clinical Trials Unit

of the Zhongshan Ophthalmic Center (Guangzhou, China) using an online random number generator (www.randomization.com) and concealed until a worker is determined eligible and has agreed to participate. The field team will have a list provided by the textile factories of potential participants, their current age, tenure and baseline productivity. Study personnel will access the random assignment for each participant according to the correct age-tenure-productivity stratum only at the time of enrolment.

4.10 Data collection schedule

4.10.1 Data collection before trial

Demographic data will be collected through Shahi's Human Resource database and clinical information will be collected using VisionSpring's Eye Examination Form (Annexe 1). Baseline efficiency data for enrolled workers will be collected from Shahi for an 8-week baseline period prior to randomization.

4.10.2 Data collection during trial

Daily efficiency data for each employee in the sewing department are collected routinely by Shahi on their Manufacturing Execution System software (SIPMONlite). Wage and employment status data will be collected on a daily basis from Shahi's Human Resource division. Employment status data will be assessed on a weekly basis over the 18-month intervention period. Intervention costs will be collected by VisionSpring and assessed at the end of the trial intervention period. Secondary outcome data will be collected using the following data collection forms

- Baseline Assessment (following eye examination)
- Endline Assessment (at the end of the 18-month intervention period)
- Spectacle Wear Compliance (on a weekly basis over the 18-month intervention period)
- Self-assessment of productivity by workers enrolled in the trial (once every 3 months stating at the baseline survey)
- Supervisor assessment of worker productivity for workers enrolled in the trial (once every 3 months stating at the baseline survey)
- Identification of trial participants leaving the employ of Shahi will occur every two months and telephone interviews will be scheduled over the course of each month

4.11 Interim analyses and stopping rules

An independent Data Monitoring and Ethics Committee (DMEC) has been established, whose remit will be to review the trial's progress. The DMEC is independent of the trial organisers. Interim analyses will be supplied, in strict confidence, to the DMEC, as frequently as the Chair requests. The DMEC Charter (Appendix 1) and Operating Procedures (Appendix 2) will be agreed before their first meeting. Meetings of the committee will be arranged periodically, as considered appropriate by the Chair. In the light of interim data on the trial's outcomes, adverse event data, accumulating evidence from other trials and any other relevant evidence, the DMEC will inform the Trial Steering Committee (TSC) if in their view there is proof beyond reasonable doubt that the data indicate that any part of the protocol under investigation is either clearly indicated or contra-indicated, either for all

participants, or for a particular subgroup of trial participants. Unless modification or cessation of the trial is recommended by the DMEC, the TSC, investigators, collaborators and administrative staff will remain ignorant of the results of the interim analysis. The accumulating trial data by arm and interim analyses will be confidential and will only be viewed by the TSC upon the recommendation of the DMEC. The TSC will not be routinely privy to these interim reports. The DMEC will make recommendations to the TSC based on the interim data. Collaborators and all others associated with the study may write to the DMEC to draw attention to any concern they may have about the possibility of harm arising from the treatment under study. The TSC Charter (Appendix 3) and its relationship to the DMEC has been discussed and agreed prior to the start of recruitment.

4.12 Trial reporting

The trial will be reported according to the principles of the CONSORT statement.

5. PROTOCOL DEVIATIONS

A protocol deviation is defined as a failure to adhere to the protocol such as the wrong intervention being administered, incorrect data being collected and documented, errors in applying inclusion/exclusion criteria or missed follow-up visits due to error.

5.1 Major

The following will be defined as major protocol deviations:

- Enrolled workers have not provided Informed Consent
- Data considered fraudulent

5.2 Minor

The following will be defined as minor protocol deviations:

5.2.1 Participants randomised in error

- Employees < 30 years, or employed < 3 months, or not employed in the sewing department assigned to Intervention Group

5.2.2 Participants who do not receive allocated intervention

- Workers in the Intervention Group not receiving glasses
- Workers in the Control Group received glasses from VisionSpring (workers in the Control Group deciding to purchase glasses from an external eyecare service provider during the course of the study will not be considered a protocol deviation, and their data will be analysed under the Intention-to-treat principle)

6. ADHERENCE TO THE INTERVENTION

Adherence to the intervention will be assessed through observation of spectacle wear while working. Adherence will be measured surreptitiously on a weekly basis. The enumerators

assessing spectacle wear compliance are GBL staff members familiar to Shahi workers. The enumerators are frequently seen working alongside Shahi workers on the production line.

7. ANALYSIS POPULATIONS

7.1 Primary analysis strategy

All outcomes will be assessed by Intention-to-treat (ITT): participants will be analysed in the groups into which they were randomly allocated, i.e. comparing the outcomes of all workers allocated to the Intervention Group with workers allocated to the Control Group, regardless of allocation received.

7.2 Descriptive analysis population

Baseline demographic and clinical characteristics will be reported for all workers randomised for whom we have data available, excluding protocol deviations randomised in error where Informed Consent has not been obtained.

8. DESCRIPTIVE ANALYSES

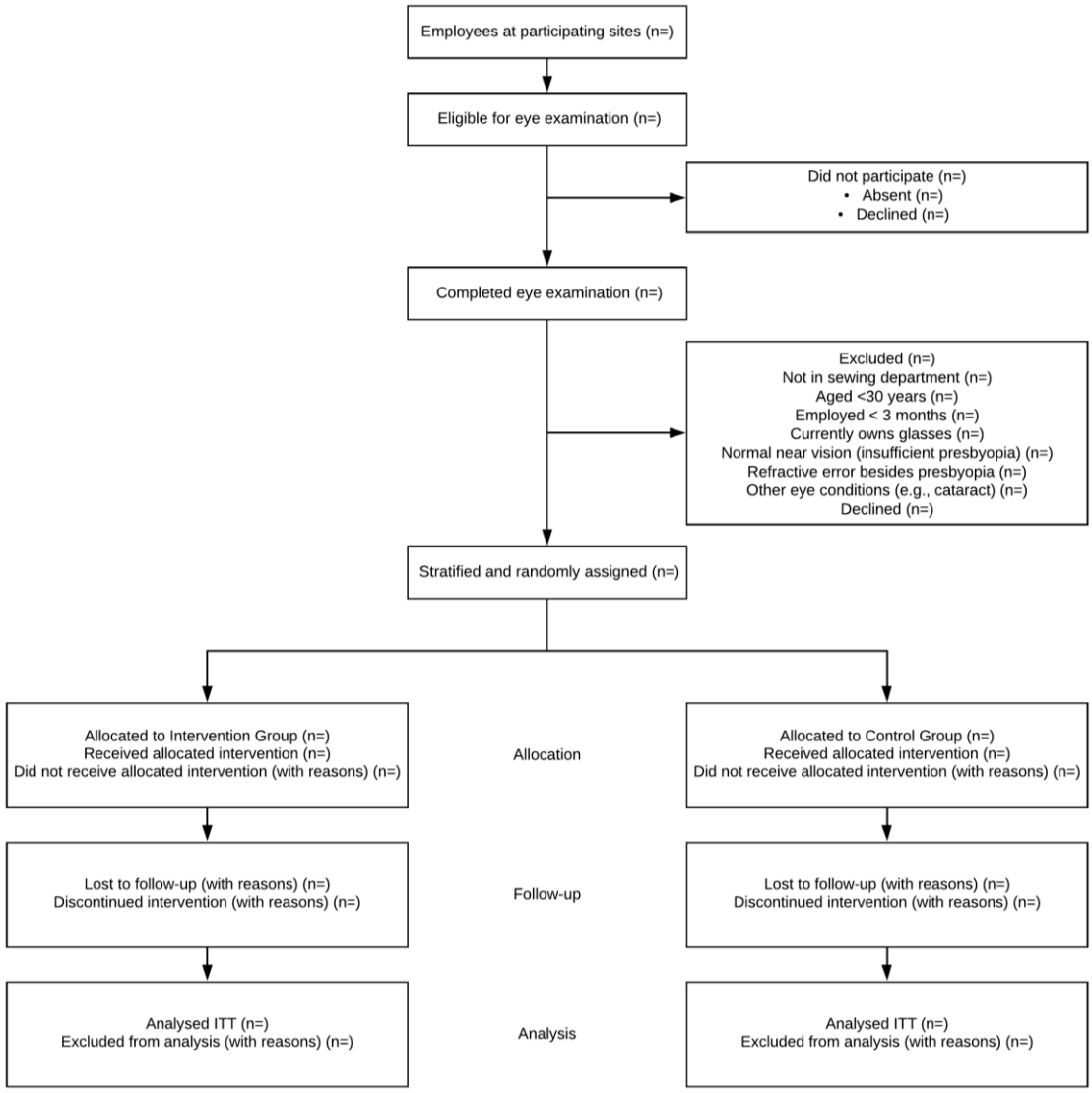
8.1 Representativeness of trial population and participant throughput

The flow of participants through each stage of the trial will be summarised using a CONSORT diagram (see below). We will report the numbers of participants:

- at participating sites
- eligible for eye examination
- examined
- stratified and randomly assigned
- received intended intervention
- withdrawals
- randomised in error

included in the analysis

CONSORT Flow Diagram



8.2 Baseline comparability of randomised groups

Participants in randomised groups will be described separately with respect to the following characteristics at trial entry:

- age
- sex
- education level
- marital status
- number of children and dependants
- median wage compared to urban Bangalore and urban India
- near vision quality of life (THRIVE tool)
- uncorrected visual acuity in each eye separately at distance and both eyes together at near
- urban or rural factory

- mean power of near correction required
- baseline work productivity
- worker self-efficacy
- attitudes towards vision correction
- access to local eyecare services
- history of uptake of eyecare services

Numbers (with percentages) for binary and categorical variables and means (and standard deviations), or medians (with lower and upper quartiles) for continuous variables will be presented; there will be no tests of statistical significance performed nor confidence intervals calculated for differences between randomised groups on any baseline variable.

8.3 Losses to follow-up

Whilst high degrees of loss-to-follow-up can lead to biased estimates of the intervention effect (particularly when there is differential drop out between intervention arms, which is related to the intervention) we anticipate minimal loss-to-follow-up with respect to the primary outcome and for secondary outcomes. In the case of the primary outcome, loss of employment at Shahi, "loss to follow-up" is in and of itself an endpoint. We will perform complete-cases analyses in each case. With respect to the covariates, we anticipate a small (<2%) amount of missing data. As the anticipated amount of missing data is small, we will analyse the data using a complete-case analysis.

8.4 Adherence to intervention

The percentage of glasses wear adherence per week will be reported for the duration of the trial.

9. COMPARATIVE ANALYSES

9.1 Analysis of primary and secondary outcomes

We will calculate the Cox proportional hazard ratio for the rate of retention between Treatment and Control groups with and without adjustment for the following co-variates: age, work tenure, spectacle wear compliance, self-reported self-efficacy, rural vs urban factory, hostel vs non hostel residence, marital and parental status, skill grade (if not too highly colinear with work tenure), work attendance. Linear regression analyses will be performed on potential determinants of primary and secondary outcomes. The study group and all significant variables with p values less than 0·20 in simple regression analyses will be included in multiple regression models. Histograms, normal quantile plots (QQ Plot), and the Shapiro-Wilk normality test will be used to test the normality assumption in regression models. For the Visual Quality of Life (THRIVE) tool, a composite score (eleven items) and near activities sub-score (five items) will be created on 0–100 scales. Intervention costs will comprise the screening test, glasses (and any replacement thereof) as well as direct and indirect costs to the company for facilitating workplace-based sight tests. We will report cost effectiveness distinguishing between study costs and program costs.

The analysis of primary and secondary outcomes will be adjusted for the minimisation factors (age, work tenure at the textile factory, efficiency during baseline assessment) to

account for the correlation between treatment groups introduced by balancing the randomisation.[5] Both the crude unadjusted and adjusted estimates will be presented, but the primary inference will be based on the adjusted analysis.

For multiple imputation of missing data in assessing primary outcomes we will create 20 copies of the data, in which missing values shall be imputed by chained equations, and the datasets will be averaged.

9.2 Pre-specified subgroup analysis

The consistency of the effect of the intervention across specific subgroups will be assessed using the statistical test of interaction. Note that this study was not designed to have sufficient power to test for interaction terms in these subgroup analyses; we will interpret the results with caution. Pre-specified subgroup analyses include:

- Sex
- Skill grade (<median, ≥median)
- Degree of Presbyopia (early presbyopia, reading glass powers < 1.25D; moderate presbyopia, 1.25D <= reading glasses powers < 2.00D, advanced presbyopia: reading glasses powers >=2.00D)
- Age group (30 to 40years, 40 to 50 years, 50 years and older)

Subgroup analysis will be performed on the primary outcome and mean efficiency outcomes.

9.3 Sensitivity analysis

We will conduct additional "post-hoc" analyses in which we will adjust for additional pre-specified potential confounders (table below). The pre-specified confounders will be included in the models even when no baseline imbalance exists. We have limited the inclusion of potential confounding variables to those that we surmise to be the most important based on the investigators' assessment of clinical plausibility. This approach has been chosen since confounder selection strategies which are based on collected data, for example selecting confounders using preliminary statistical tests, result in models with poor statistical properties such as incorrect type I error rates.[6-9] Those confounders that are highlighted as having a significant impact in the subgroup analyses will also be included in these analyses to assess the impact on the treatment effect.

Level	No.	Confounder
Patient	1	Age (years, continuous)
	2	Sex (Male/Female)
Site	3	Rural/Urban

9.4 Significance levels and adjustment of p-values for multiplicity

For the primary and secondary outcome, including subgroup analyses, a 95% confidence interval will be calculated.

9.5 Procedure for accounting for missing, unused, and spurious data

Missing data will be described, for example, by presenting the number and percentage of individuals in the missing category. All data collected on data collection forms will be used, since only essential data items will be collected.

9.6 Exclusion of data

Before data are locked for statistical analysis, a blinded review of all data will take place. Any decision to exclude a subject or single observation from the statistical analysis is the joint responsibility of the PROSPER III trial statistician, the international trial manager and the Chief Investigator. Exclusion of data from analyses will be used restrictively and normally no data should be excluded from the full analysis set. The subjects or observations to be excluded, and the reasons for their exclusion will be documented and signed by those responsible before database lock. The subjects and observations excluded from analysis sets, and the reason for this, will be described in the clinical trial report. Any observation excluded from the analysis database will be documented before database lock with the reason for exclusion provided.

9.7 Statistical software employed

Stata, version 16.1

10. QUALITATIVE SUB-STUDY

One of the hypotheses being tested in PROSPER III is providing glasses to correct presbyopia in sew-ers in India can reduce attrition. The RCT will collect data comparing the attrition rates of Intervention and Control group subjects; however, this will not explain why sew-ers have left their jobs, and the role that vision plays in that decision. This qualitative sub-study will provide this additional information.

10.1 Research Design

The descriptive qualitative study design is frequently used in mixed methods research because its outcome – a straightforward report on methods and findings close to participants' own voices – makes it relatively easy to integrate quantitative and qualitative findings. Furthermore, our aim is to highlight the voice of sew-ers who have left Shahi during the study rather than develop a theoretical model or explore pre-determined factors thought to affect their decisions.

10.2 Sample and Sample Size

The sampling frame or set of potential participants will be all PROSPER III subjects who leave the employ of Shahi before the end of the study. In a descriptive qualitative study, approximately 20 participants can provide enough data to reach informational redundancy or data saturation (the point when no new information emerges from the last round of data collected) and theoretical saturation (the point when the researchers identify themes or categories with enough data to support their interpretations of each one). If some participants do not provide rich data, then additional persons must be recruited to reach

saturation. Because we will be selecting participants from both the Intervention and Control groups, we will increase the sample size to 40 persons, planning for an additional five interviews if needed. Administrative data from Shahi indicate that about 20% of their workers leave after 5 months of employment, 50% after 12 months and 60% after 18 months. Assuming that the attrition rate for sew-ers is similar, we will have a large enough sampling frame to identify and enroll 40-45 participants.

10.3 Recruitment

Using administrative data at the enrolled factories, we will identify job leavers once every two months, compare the lists with the enrollment list for PROSPER III, and create a random-ordered list of job leavers to invite into this sub-study. A research team member will attempt to contact each person via telephone on the list in sequence, making up to three calls before declaring them 'unreachable' and removing their name from the recruitment list. We will provide interested sew-ers an information and consent form specifically related to the qualitative sub-study, instructing them to take at least 48 hours to consider whether they want to join this part of the study. The researcher will then call each contacted person again after the 48-hour period to verify her interest in being interviewed.

10.4 Time Period

Prior work in the area has revealed that factory workers often change their mobile phone number without notice, making it more difficult to find and recruit participants for this sub-study. Therefore, we will ensure we have enough time to enroll 40-45 participants by making recruitment calls every other month starting at month six of PROSPER III, continuing until we have completed enough interviews to reach saturation.

10.5 Data Collection

We will use a standardized protocol to back-translate the relevant documents (sub-study information sheet, consent form and interview guide). Using the first draft version of the instruments, we will pilot test them for acceptability and clarity with three to five participants in each of the study languages. We will use feedback from the pilot test participants to revise the instruments and test them again with one person for each language, making final adjustments as needed. Seven to 10 days after recruiting a participant, the interviewer will call to: a) review the sub-study information sheet, b) ensure the participant still wants to be interviewed, and c) conduct the phone interview or schedule it for another time. The interviewer will use the semi-structured interview guide and audiotape (with prior permission) the oral consent and interview. S/he will also audiotape a post-interview reflection capturing information about the quality and likely veracity of the interview, unanticipated information that may be incorporated into an updated version of the interview guide, and other observations.

10.6 Data Analysis

The audiotaped interviews will be transcribed in the original language (Hindi, Kannada or Odia) by the interviewer, who will ensure accuracy by comparing the audio and written versions of the interview. A bilingual translator will produce an English-language version of

each transcript, conferring with the interviewer to ensure accuracy. The analysts will carry out the following steps:

10.6.1 Prepare transcripts.

They will make a working copy of each transcript, add line numbers (to help contextualise quotes during analysis), set aside less relevant statements and labeling each meaning unit (quote) with an anonymous source code (e.g. 'Int-6' = 6th participant from the Intervention Group).

10.6.2 Produce analysis matrices

The analysts will create a blank table for each topic from the interview guide that they will populate with relevant quotes from transcripts (one quote per row). The tables will consist of four columns, one for the row number, another for the quote, then one for a brief summary of the quote's meaning and finally one for a code/sub-code to be developed from the meaning statements.

10.6.3 Label (code) data

Starting with one transcript, two analysts will independently populate the tables and write a brief statement about each quote. They will confer and reach consensus on this work and then independently assign labels (codes and sub-codes) based on the meaning statements. They will confer again, reaching consensus on the codes and sub-codes. The aim is to reduce the number of labels used to 5-8 codes per topic.

10.6.4 Create the list of codes and sub-codes

Each time the analysts meet to discuss their codes, they will update a document with the list of codes and sub-codes, each with a definition and exemplary quote. They will also update the document, noting the new version code, date and source transcripts.

10.6.5 Cluster data using the codes/sub-codes

The next step is to combine all the data from the individual tables for a specific topic into a single master table. The analysts can then discuss codes that have a large number of entries, deciding if they should create new sub-codes or codes and split the data. Likewise, they will discuss codes and sub-codes that are rarely used, either labelling them as unique or unanticipated findings or folding them into another code or sub-code.

10.6.6 Sort data by salience and relevance

The analysts will examine the codes used for a particular topic and list them in rank order both for the number of times each code was used (salience) and the proportion of participants whose statements are represented by the a code (relevance). The analysts will present the rank-ordered results in tabular form based on the number and proportion of mentions, noting the associated proportion of participants. Responses (codes) that are both highly salient (>80% of the mentions) and highly relevant (noted by >80% of the

participants) will be the focus of discussion about domains or topics. They will repeat the process, identifying the most salient and relevant sub-codes for each code.

10.6.7 Thematic analysis

Another important way to present findings wil be for the analysts to review the entire dataset and identify one or two overarching themes at a 'macro level' of analysis.

10.6.8 Write report

The analysts will prepare a report on methods used to collect and analyse the data, as well as findings from both the qualitative content analysis (salience and relevance of categories within each topic) and thematic analysis.

10.7 Outcomes

- A rank-ordered list of reasons sew-ers have left Shahi or the textile industry
- A report on the role that poor near vision (uncorrected presbyopia) has on sew-ers' productivity, self-efficacy and job leaving
- Material for wider dissemination to academic audiences through presentations at conferences and publications in peer-reviewed journals.

10.8 Rigour

We will take several steps to ensure the qualitative sub-study is rigorous by following standard procedures aimed to fulfil the four criteria of excellence for qualitative research.

10.8.1 Credibility

Credibility (equivalent to internal validity) will be ensured through: a) prolonged and varied engagement with participants throughout the entire PROSPER III study; b) seeking referential adequacy by collecting documents and field notes that describe the study context; c) triangulation (the use of multiple data types and sources regarding worker attrition, and multiple investigators and perspectives); and d) producing high-quality findings through following a well-designed protocol for the sub-study, training the investigators, having an experienced qualitative researcher analyse field notes and analytic memos, and peer debriefing about the procedures and findings with qualitative experts not involved with the study.

10.8.2 Transferability

Transferability (equivalent to external validity or generalizability) will be enhanced by a) providing a rich description of methods, context and findings; b) ensuring that sampling to data saturation was achieved; and c) data collection and analysis are done in a transparent way. This will allow others to decide if the methods and findings could be transferred to their setting or population (case-to-case generalization) or added to the literature on the study subject.

10.8.3 Dependability

Dependability (findings would be similar if the same cohort of participants, analysts and contexts existed) will be developed by maintaining a rich description of study methods and an audit trail or copy of all documents generated during the sub-study to allow others to inspect the methods used to reach study concisions. Other methods include stepwise replication of the data analysis procedures through assessing coding accuracy and inter-coder reliability.

10.8.4 Confirmability

Confirmability (confidence that results can be corroborated by other researchers with similar participants, methods and instruments) will be enhanced by triangulation (explained above), presenting information on the researchers and their positionality (e.g., experience with the research topic or study participants). Researchers will keep a 'reflexive journal' and share their experiences with the team at regular meetings, ensuring that their perspective does not overshadow the views and experiences of participants.

11. SAFETY DATA ANALYSIS

11.1 Serious adverse events

Serious adverse events will be listed by allocation as well as allocation received.

12. ADDITIONAL EXPLORATORY ANALYSIS

Any analyses not specified in the analysis protocol will be exploratory in nature and a 2-sided significance level of 0.01 will be used with 99% confidence intervals.

13. DEVIATION FORM ANALYSIS DESCRIBED IN PROTOCOL

None at present

13. REFERENCES

1. Fricke TR, Tahhan N, Resnikoff S, et al. Global Prevalence of Presbyopia and Vision Impairment from Uncorrected Presbyopia. Ophthalmology. 2018;125(10): 1492—1499.
2. Reddy P, Congdon N, MacKenzie GE, et al. Effect of providing near glasses on productivity among rural Indian tea workers with presbyopia (PROSPER): a randomized trial. Lancet Glob Health. 2018;6: e1019—1027.
3. Li R, Chen X, Yan H, et al. Functional consequences of iron supplementation in iron-deficient female cotton mill workers in Beijing, China. Am J Clin Nutr. 1994;59: 908—913.
4. Wolgemuth JC, Latham MC, Hall A, Chesker A, Crompton DWT. Worker productivity and the nutritional status of Kenyan road construction laborers. Am J Clin Nutr 1982; 36: 68—78.
5. Basta SS, Soekirman, Karyadi D, Scrimshaw NS. Iron deficiency anemia and the productivity of adult males in Indonesia. Am J Clin Nutr 1979; 32: 916—925.5:e888.
6. Kahan BC, Morris TP. Reporting and analysis of trials using stratified randomisation in leading medical journals: review and reanalysis. BMJ 2012;345:e5840.
7. Permutt T. Testing for imbalance of covariates in controlled experiments. Stat Med 1990; 9(12):1455-1462/
8. Pocock SJ, Assmann SE, Enos LE, Kasten LE. Subgroup analysis, covariate adjustment and baseline comparisons in clinical trial reporting: current practice and problems. Stat Med 2022; 21(19):2917-2930.
9. Senn SJ. Covariate imbalance and random allocation in clinical trials. Stat Med 1989; 8(4):467-475.

14. APPROVAL

Senior Statistician	Jin Ling	
Chief Investigator	Nathan Congdon	