Pre-Analysis Plan

Promotion of improved cookstoves by the Red Cross in rural Ethiopia as part of the 3FM pneumonia project.

How effective are subsidies and information provision?

MSc Thesis Development Economics (DEC-80436)
Wageningen University

Joy de Korte
Supervisor: Maarten Voors

March 2019

Acknowledgements

You are about to read my MSc thesis, a document on which I have been working during the last ten months. Since my assignment was to write a pre-analysis plan, the writing of this thesis has not exactly been a linear process. We have been developing new ideas and changing our plans continuously over the last few months. Of course, a lot of people were involved in this challenging as well as inspiring process. I would like to take the opportunity to thank these amazing people for their contribution to my thesis.

I am very thankful for the chance given to me by the Netherlands Red Cross and Ethiopian Red Cross societies to work on this study. I would like to thank Lotte and Jette for introducing me to the world of the Red Cross and supporting me in the process of writing this thesis. I will definitely miss our weekly coffee meetings. Another special thanks goes to the NLRC team in Ethiopia for their valuable cooperation and assistance. You made my stay in Ethiopia an unforgettable experience. Thank you for teaching me how to eat *injera* and all the other things that I needed to learn in order to survive in your beautiful country.

Furthermore, I would like to thank my supervisor, Maarten Voors, for his patience and enthusiasm. Your passion for research is truly inspiring. Thanks also to Jörg Peters for sharing his knowledge and love for improved cooking stoves.

To conclude, a big shout out to my beloved family and friends. You always motivate me to follow my dreams and get the best out of myself. Thanks for all your practical support, motivational speeches and for listening to my endless monologues about improved cooking stoves, randomized treatments and statistical power.

Abstract

This document provides a pre-analysis plan for the evaluation of an improved cooking stove intervention. This intervention will be implemented by the Ethiopian and Netherlands Red Cross Societies as part of the 3FM pneumonia project. The goal of this project is to increase community resilience by reducing the mortality and morbidity rates of acute respiratory infections among under-fives. We expect to conduct the impact evaluation in a total of 20 sub-districts of the South Gondar and Guji zones in Ethiopia.

By means of a Randomized Controlled Trial, we aim at studying the impact of subsidies and information provision on *Mirt* stove adoption and usage as well as economic and health outcomes. With the provision of information on *Mirt* stove benefits, we aim at influencing the intra-household bargaining processes linked to stove adoption. Because of a mismatch between stove (choice) preference and intra-household bargaining power, it is crucial to motivate men to support their wives in this adoption decision. Since women generally lack the authority to make the stove purchase decision. We will vary the information provision on *Mirt* stove benefits in order to assess if awareness on specific stove benefits leads to increased ICS adoption, through intra-household bargaining processes. Furthermore, we will offer stove subsidies with the aim to reduce the liquidity- and credit constraints of rural Ethiopian households. Different levels of subsidies will be introduced to learn more about the relationship between the level of stove subsidy and ICS adoption.

With these two treatment factors, eight different treatment conditions can be created. These treatment conditions will be randomized on village level. Approximately 483 villages will be included in the randomization process.

This document includes an overview of the baseline situation in the study areas, the design of the impact study and an analysis of experts' predictions regarding the results of this study.

Keywords: *Pre-Analysis Plan, Randomized Controlled Trial, Improved Cooking Stoves, subsidy, intra-household bargaining process, stove adoption, Mirt stove, Ethiopia, International Red Cross and Red Crescent Movement*

List of tables

List of Figures

Abbreviations

ARI	Acute Respiratory Infection
BCC	Behavioural Change Communication
CBHFA	Community Based Health and First Aid
CHERG	Child Health Epidemiology Reference Group
CO	Carbon Monoxide
CSA	Central Statistical Agency
ERCS	Ethiopian Red Cross Society
FGD	Focus Group Discussion
GIZ	Deutsche Gesellschaft für Internationale Zusammenarbeit
HEW	Health Extension Worker
IAP	Indoor Air Pollution
IAQ	Indoor Air Quality
ICC	Intraclass Correlation Coefficient
ICS	Improved Cooking Stove
ITT	Intent-To-Treat
MDE	Minimal Detectable Effect
NGO	Non-Governmental Organisation
NLRC	Netherlands Red Cross Society
PAP	Pre-Analysis Plan
PfR	Partners for Resilience
PM	Particulate Matter
$PM_{2.5}$	Particulate Matter (smaller than 2,5 µm)
RC	Red Cross
RCT	Randomised Controlled Trial
SDU	Standard Deviation Unit
TOT	Treatment-On-the-Treated
VIF	Variance Inflation Factor
WHO	World Health Organization
WTP	Willingness To Pay
WUR	Wageningen University & Research
ZRCS	Zambia Red Cross Society

Table of contents

1. Introduction

Pneumonia is one of the leading causes of under-five mortality in developing countries. Likewise, in Ethiopia, pneumonia is the leading single disease killing children under five years of age (UNICEF, 2014). According to estimates of the Child Health Epidemiology Reference Group (CHERG) and the World Health Organisation (WHO) from 2014, 3.4 million Ethiopian children encounter pneumonia annually. At the same time, coverage of proper case management of Acute Respiratory Infections (ARI), such as pneumonia, has been worryingly low in the country. Only 30 percent of the caretakers indicated to have sought care from a health-care provider when a child below five years of age experienced symptoms of ARI (Central Statistical Agency of Ethiopia [CSA] & ICF International, 2016).

The Netherlands Red Cross Society (NLRC) recognizes pneumonia as an imminent threat to child health and decided that the proceeds of the yearly radio fundraising event "3FM Serious Request" from 2016 would be spend on a project that focuses on pneumonia reduction. A total of 9.25 million euros was raised for the implementation of this project. These funds have been used for the implementation of the 3FM pneumonia project (2017-2020) in five different countries, including Ethiopia. In Ethiopia, the project is implemented in the Amhara, Somali and Oromia regions. The overall objective of this pneumonia project is to increase community resilience by reducing the mortality and morbidity rates of ARI among under-fives in Ethiopia. This multi-faceted intervention aims at addressing various risk factors for pneumonia in children younger than five.

The potential significant risk factors for developing pneumonia in under-fives have been studied extensively during the last three decades. Strong evidence was found for the role that child's nutritional and vaccination status as well as other childhood infectious diseases, such as measles and diarrhoea, play in the development of childhood pneumonia. Overall, children with compromised immune systems, children born with very low birth weight as well as malnourished children are at higher risk of developing pneumonia. In addition, limited access to medical care is identified as risk factor for developing pneumonia in under-fives. Furthermore, crowded homes and exposure to Indoor Air Pollution (IAP) are recognised as environmental factors that increase children's susceptibility to pneumonia (Geleta, Tessema, & Ewnetu, 2016; Lanata et al., 2004).

Traditional biomass-fired cooking stoves contribute substantially to IAP through their high emissions of health-damaging pollutants such as carbon monoxide (CO) and fine particulate matter ($PM_{2.5}$). This is caused by the typical construction of these traditional cooking stoves which results in incomplete and inefficient combustion (Panwar, Kurchania, & Rathore, 2009; Smith et al., 2000). Women and young children suffer most from the substantial emissions of air pollutants of these cooking stoves, as they often spend long periods of time in close proximity to indoor fires during cooking activities (Barnes, Mathee, & Moiloa, 2005; Bruce et al., 2006; Ezzati & Kammen, 2000). Continuous inhalation of particulate matter emitted by traditional cookstoves is strongly linked to the development of pneumonia (Barnes, 2014; Fullerton, Bruce, & Gordon, 2008; McCracken et al., 2007). Exposure to IAP has an exceptionally strong harmful effect on children's respiratory health, since their immune system as well as the epithelial linings of their lungs are not yet fully developed. This limits their body's defence against infections and leads to a greater permeability of pollutants (Pande, 2000; Smith et al., 2000). Furthermore, they have relatively narrow airways and a relatively large lung surface area per kilogram of body weight. As a result of the latter, they breath in approximately 50 percent more (polluted) air under normal breathing conditions than adults (Moya, Bearer, & Etzal, 2004).

Taking into account the negative health effects of traditional cookstoves, the Red Cross (RC) decided that the 3FM pneumonia project should, amongst others, address harmful cooking practices. Today, Ethiopian households still rely predominantly on traditional stoves and biomass resources to prepare their meals. Nearly all rural Ethiopian households (98.4%) exclusively cook on the so-called traditional three-stone stove (CSA, 2012). The charcoal stove is the second most commonly used stove in Ethiopia (Gaia association, 2014). As part of the 3FM pneumonia project, an Improved Cooking Stove (ICS) intervention will be implemented to promote improved biomass stoves in rural areas of Ethiopia.

According to the World Health Organisation (WHO, 2018), improved cookstove technologies can be quite efficacious in reducing the public health burden. This claim specifically concerns smoke-free ICSs that comply with the WHO air quality guidelines. Other organisations and researchers also refer to simpler, fuel saving, non-traditional biomass stoves as being ICSs. These stoves are not necessarily clean according to the WHO standards. The term 'improved cooking stove' thus describes a wide range of replacements for the traditional cooking method. A large variation is observed in the degree to which these so-called improved stoves actually burn cleaner (Jetter et al., 2012).

Nevertheless, ICS usage is generally reported to reduce smoke production and emission levels of air pollutants (Bensch & Peters, 2015; Chowdhury et al., 2012; Seguin, Flax, & Jagger, 2018). As a result, ICS usage can improve respiratory health of women and children significantly (Bensch & Peters, 2015; Jamali et al., 2017; Schilmann et al., 2015). Furthermore, these more efficient stoves reduce fuel consumption. This results in a reduction in workload, time-losses and health impacts, or monetary expenses associated with fuelwood collection as well as pressures on natural forests (Bailis et al., 2015; Parikh, 2011).

However, the realisation of the benefits linked to ICS usage has proved to be quite challenging. For instance, multiple barriers prevent households from switching to ICSs in the first place. These barriers include costs, infrastructure, technology, information and socio-cultural factors (Schlag & Zuzarte, 2008; Rehfuess et al., 2013). Despite the efforts of the Ethiopian government and the private sector to boost adoption of fuel-saving cooking stoves, only few rural Ethiopian households own a non-traditional stove. Furthermore, stove adoption alone is not sufficient. Long-term exclusive use of ICSs is crucial in order to realise the above-discussed benefits (Clark, Heiderscheidt, & Peel, 2015; Johnson & Chiang, 2015).

With these barriers to adoption of ICSs in mind, Wageningen University (WUR) developed a study to evaluate the ICS intervention that will be implemented as part of the 3FM pneumonia project. The aim of this study is to gain greater understanding of the adoption process of ICSs and the impact of ICS usage. As mentioned before, it is quite a challenge to increase stove adoption among rural households, since multiple barriers prevent these households from switching to ICSs. Amongst others, financial constraints and low levels of awareness on benefits of adoption are found to prevent households from adopting an ICS. Furthermore, a mismatch exists between stove (choice) preference and intra-household bargaining power. Women generally tend to have stronger preference for ICSs than men, since they bear the largest share of the health and other burdens associated with the use of traditional stoves. However, they generally lack the authority to make the actual stove purchase decision (Miller & Mobarak, 2013). Therefore, it is crucial to convince men of the benefits of ICS adoption in order to motivate them to support their wives in this adoption decision.

By means of a Randomized Controlled Trial (RCT), we aim at studying the extent to which subsidies and information provision on stove benefits affect ICS adoption and usage as well as economic and health outcomes.

This study will thus contribute to the literature on household adoption behaviour towards ICSs. It will provide deeper insights into the extent to which subsidies affect the adoption of improved stoves. In addition, the outcomes of this study will inform us on the potential of information provision on stove benefits as a strategy to influence the intra-household bargaining process related to ICS adoption.

From a policy perspective, the results of this evaluation will contribute to effective strategies for maximising the adoption of improved stoves. The RC can also use these results to inform the design of future projects that promote other new technologies.

This master's thesis focuses on the context and design of the study and thereby answers the following exploratory research questions:

I. Which risk factors are associated with the development of childhood pneumonia in the project areas?

II. Which strategy is adopted by the RC to reduce pneumonia among under-fives in Ethiopia?

III. How effective are stove subsidies and information provision on stove benefits in increasing ICS adoption?

IV. What are experts' beliefs on the impact of this ICS intervention and how do these predictions compare to results from previous studies on comparable interventions?

In the remainder of this document I answer these research questions. The following section summarises the existing literature on risk factors of developing pneumonia in under-fives, stove adoption and barriers to adoption. After which, section 3 analyses respiratory child health and the related risk factors in the project area. Next, section 4 provides more details on the 3FM pneumonia project and section 5 describes the design of the impact study. Expert predictions regarding the study outcomes are analysed in section 6, while section 7 wraps up with some concluding remarks.

This document provides a Pre-Analysis Plan (PAP) for the evaluation of an ICS intervention that will be implemented by the RC. It's a step-by-step plan setting out how we will conduct this study and collect and analyse our data. Only the design of the study will be covered. This document does not contain any study results. Recently the demand for PAPs has increased in social sciences. It asks researchers to pre-specify the analysis to be run before collecting and examining the data, which requires them to report all pre-specified results. As a result, data mining becomes much less of a problem and trial transparency improves. Furthermore, the writing of PAPs may lead to improvements in the quality of trials by making it possible to identify potential problems early in the research process. Since the PAP provides a roadmap for the data analysis, collected data can be analysed much quicker and easier as well.

Study designs can be registered in a social science registry, such as the EGA, RIDIE or AEA RCT registry. Registration of interventional trials is considered to be a scientific and ethical responsibility since it makes it easier for program designers and policymakers to access all available evidence. This counteracts the problems of publication bias and selective reporting and enables them to make informed decisions.

2. Literature review

This section starts with a brief overview of the literature on risk factors for childhood pneumonia. Subsequently, the way in which the ICS intervention aims at reducing the burden of childhood pneumonia is clarified once again. The section concludes with an overview of barriers to ICS adoption and ways in which these barriers can be reduced.

2.1 Risk factors for pneumonia

As earlier discussed, numerous evaluations of risk factors for childhood pneumonia in developing countries have been conducted. This sub-section provides an overview of the risk factors that were found to be consistently associated with childhood pneumonia. The approach adopted to construct a list of established risk factors for childhood pneumonia, is based on the procedure applied in the Pneumonia Etiology Research for Child Health Study (Wonodi et al., 2012).

The list of prioritised risk factors for developing pneumonia in under-fives from the CHERG provided a good starting point for an overview of the established risk factors for childhood pneumonia. Subsequently, this set of significant risk factors was substantiated as well as augmented by an extensive literature review. The resulting list of risk factors is divided in the following categories: child demographics, birth details, nutrition, past morbidities or comorbidities, child's vaccination status, maternal characteristics, family environment and local health care system. Table 1 presents the final (non-exhaustive) list of established risk factors for developing pneumonia in under-fives.

These risk factors have been targeted by various types of interventions during the last three decades. The ICS intervention that will be evaluated by means of this impact study, aims to reduce the burden of childhood pneumonia through reducing *Exposure to indoor air pollution* (See Table 1). The promotion and distribution of ICSs has become an important and integral component of development initiatives and is supported by programs, governments and donor agencies all over the world. Preferably, a more holistic approach would be adopted by the RC to change cooking practices and reduce IAP. However, in order to assess the impact of this technical intervention, it is essential to implement this stove intervention independently of other activities. Therefore, we will implement and evaluate the stove intervention in isolation.

Table 1. *Risk factors for developing pneumonia in under-fives.*

Risk factor category	Risk factor variable	Source
Child demographics	Age	1
Birth details	Birth weight	2, 3, 4, 5, 9
Nutrition	Breastfeeding practices	2, 4, 6, 7, 18
	Malnutrition (or undernutrition)	2, 3, 4, 5, 6, 8, 18, 19, 20
	Zinc deficiency	6, 22, 23
	Vitamin D deficiency	20, 24, 25
Past morbidities or comorbidities	Diarrhoea	10
	Measles	6, 11
	AIDS/HIV infections	11, 12, 26
	Malaria	12
	History of pneumonia or wheezing	8, 5
Child's vaccination status	Measles	13
	Pertussis	14
	Hib	15
	Child immunization coverage	3, 4, 7, 8, 27
Maternal characteristics	Maternal age	5, 16
	Maternal education	17, 28
Family environment	Crowding level	2, 3, 4, 5, 6, 8, 11, 29
	Parental smoking	9, 11, 30, 31
	Exposure to indoor air pollution	4, 6, 11, 32, 33
	Upper respiratory infection in mother or siblings	7
Local health care system	Description of local health care system (quality of care, availability of treatments)	21
	Access to health care (distance, time and costs to nearest HF)	21

Note*:* This table summarises the risk factors for childhood pneumonia in developing countries in the literature. Column 1 specifies the eight categories of risk factors, column 2 contains the individual risk factors and column 3 refers to a selection of the authors that have studied the specific risk factors. For details on these authors see below.

Sources: [1]Shah, Ramankutty, Premila & Sathy (1994); [2]Rudan et al. (2013); [3]Coles et al. (2005); [4]Rudan, Boschi-Pinto, Biloglav, Mulholland & Campbell (2008); [5]Victora, Fuchs, Flores, Fonseca & Kirkwood (1994); [6]Goetghebuer, Kwiatkowski, Thomson & Hull (2004); [7]Walker et al. (2013); [8]Broor et al. (2001); [9]Grant et al. (2012); [10]Fatmi & White (2002); [11]Hassan & Al-Sadoon (2001); [12]Muhe, Lulseged, Mason & Simoes (1997); [13]Osendarp et al. (2002); [14]Baqui et al. (2003); [15]Wayse, Yousafzai, Mogale & Filteau (2004); [16]Leis et al. (2012); [17]Schmidt, Cairncross, Barreto, Clasen & Genser (2009); [18]World Health Organization (2016); [19]Black, Brown & Becker (1982); [20]Koyanagi et al. (2011); [21]Oyejide & Osinusi (1990); [22]Mulholland (1995); [23]Mulholland et al. (1997); [24] Savitha et al. (2007); [25]Selwyn (1990); [26]Wonodi et al. (2012); [27]Macedo et al. (2007); [28]Cardoso et al. (2004); [29]Armstrong & Campbell (1991); [30]Suzuki et al. (2009); [31]Murray et al. (2012); [32]Smith et al. (2011); [33]Bang et al. (1990)

2.2 Adoption of ICSs

Experience has shown that securing adoption of ICSs can be very challenging for reasons that involve a wide range of factors (Rehfuess, et al., 2013). Over the last thirty years, ICSs have been distributed as an alternative household fuel burning device all over the world (Chowdhury et al., 2011). Despite their health, environmental and economic benefits, the adoption rates of ICSs remain worryingly low. It is especially challenging to increase adoption of non-traditional cooking technologies in rural communities where the awareness on IAP and improved cooking technologies is relatively low (Nguyen, 2017).

Barriers and facilitators to ICS adoption

A broad range of possible explanations for low adoption rates of new cost-effective technologies in developing countries have been discussed in the literature. Amongst others, liquidity- or credit constraints, low levels of awareness on benefits of adoption, neglect of externalities and lack of incentives for innovation were found to prevent households from adopting such new technologies (Conley & Udry, 2010; Dupas & Robinson, 2013; Giné & Yang, 2009; Kremer & Miguel, 2007; Tarozzi et al., 2011).

Various studies have confirmed the important role that economic variables, such as product price and household income, play in ICS adoption. Affordability has proven to be an important prerequisite for the adoption of an ICS. Since most households are credit constrained and chronically short on cash, the adoption decision is highly responsive to price (Arif et al., 2011; Bensch, Grimm, & Peters, 2015; Lewis & Pattanayak, 2012; Mobarak et al., 2012). Generally, full-cost pricing results in extremely low adoption rates.

In addition, households' decision to adopt an ICS is heavily influenced by intra-household interactions and social learning processes. Looking at the intra-household bargaining process, Miller and Mobarak (2013) have found that women – who bear disproportionate cooking costs – tend to have stronger preference for ICSs than men. Unfortunately, they often lack the authority to act upon these preferences and make the actual purchase. Men, on the other hand, will only support their wives in the adoption of an ICS if they are convinced of the benefits of this investment (Pattanayak & Pfaff, 2009). The low intra-household bargaining power of women is therefore also perceived to be an important barrier to ICS adoption.

Furthermore, households' decision to adopt an ICS is significantly influenced by opinion leaders and social networks. By observing early stove adopters, people can be motivated to adopt an ICS as well. Such an observational learning process among people is also referred to as "social learning" or "peer learning". Highly cohesive communities and well-connected early stove adopters are found to facilitate quick stove adoption (Geary, Prabawanti, & Aristani, 2014). Miller & Mobarak (2014) found the role of social learning processes in stove adoption to be particularly important when the new technology has less evident benefits or when information on the new technology is limited. Over time, when common experience accumulates, the value of information acquired from opinion leaders declines (Miller & Mobarak, 2014). The potential facilitating role of opinion leaders in stove adoption thus slowly disappears with time.

Reducing barriers to adoption

Development actors have been promoting ICSs since the 1970s. Evaluations of various technology-based interventions have resulted in a deeper understanding of their effectiveness in promoting these innovative technologies. The main finding of these evaluations is that an ICS intervention should tackle multiple barriers to adoption in order to substantially increase ICS adoption. An intervention existing solely out of a subsidy is, in most cases, not very effective in increasing the long-term adoption and usage of new technologies (Hueso & Bell, 2013). On the other hand, awareness raising alone is also not expected to substantially increase coverage of the specific technology (Guiteras, Levinsohn, & Mobarak, 2015). In line with these statements, Mobarak et al. (2012) argue that it is essential to relax households' financial constraints and raise awareness on benefits of ICSs at the same time.

By raising awareness on the benefits of ICSs and addressing perceptions and expectations on the impact of ICSs, demand for improved stoves is created. The suitability of demand creating strategies depends strongly on the context. Such a strategy should comprise both general awareness raising activities as well as person-to-person interaction (Rehfuess et al., 2013). In addition, a successful demand creating strategy should address the gender differences in preferences and decision-making power (Miller & Mobarak, 2013). A promising approach may be to promote attributes of the new cook stove technology that men value and cannot easily unbundle.

Besides creating demand for ICSs, households' financial barriers should be reduced in order to enable them to actually purchase the stove. Subsidy interventions are particularly cost-effective in reducing financial barriers (Cohen & Dupas, 2010; Dupas, 2009). However, subsidies are also quite controversial, since practitioners are concerned that they may undermine intrinsic motivation or cause dependency (Kar & Pasteur, 2005). The long-term effects of distributing subsidised stoves on stove adoption and usage are therefore heavily debated.

Stove subsidies might negatively affect future willingness to pay (WTP) if people anchor their valuations on subsidized prices (Kőszegi & Rabin, 2006). Bensch and Peters (2017) have conducted an experiment to assess these potential anchoring effects. Their outcomes indicate that a one-time free distribution does not necessarily affect the WTP for improved biomass cookstoves negatively in the long run.

Another concern is that subsidies might reduce the potential of hypothesized screening and sunk cost effects (Ashraf, Berry, & Shapiro, 2010). Subsidies could undermine the screening effect of prices if subsidized technology ends up at those who do not value, and therefore not use, it. The sunk cost effect refers to the empirical finding that people feel more compelled to put a product to good use when they have paid for it. Findings of Bluffstone et al. (2017), contradict with the existence of this psychological effect. Their findings suggest that actual ICS usage does not depend on monetary payments made. They even argue that distributing stoves free of charge is the preferred policy for promoting sustained stove usage.

All in all, various researchers have made a case for the implementation of direct promotion strategies, such as stove subsidies or even free distribution, to reduce liquidity constraints and overcome under adoption (Bensch et al., 2015; Mobarak et al., 2012). These recommendations are supported by the hypothesised learning effects. Since it has been argued that subsidies enable households to test a new technology and learn about its benefits, stove subsidies can also have a positive effect on (future) WTP.

When barriers to household adoption of ICS are successfully reduced, ICS adoption rates are expected to increase substantially. Social influence can then give rise to a self-perpetuating cycle in which early adopters stimulate even more people to adopt. The social multiplier in demand varies heavily between contexts and is strongly related to the saturation of the ICS. For example, Guiteras et al. (2015) have assessed the effects of a targeted subsidy on sanitation coverage. They indeed found that over time adoption also increased among people that were ineligible for the subsidy (Guiteras et al., 2015). In the context of the study of Guiteras et al. (2015), the move from subsidizing 25% to subsidizing 50% of the poorest households was found to produce the largest demand spillovers.

3. Child pneumonia and associated risk factors in Ethiopia

This section presents the results for RQ1: *Which risk factors are associated with the development of childhood pneumonia in the project areas?*. This research question is answered by the use of baseline data that was collected for the 3FM pneumonia project as a whole. The baseline survey was conducted by the Ethiopian Red Cross Society (ERCS) and NLRC in July 2017. Furthermore, the list of risk factors for developing childhood pneumonia, discussed in the previous section, will be used for this analysis. The results of this analysis will indicate which of these risk factors are significantly associated with the prevalence of pneumonia among the under-fives who are included in the baseline dataset.

This section first discusses the methodology of the analysis of the baseline dataset and then presents the empirical results. Subsequently, the results as well as the limitations of this analysis are discussed.

3.1 Methodology

Baseline sample

The RC used an online sample size calculator to calculate the sample size needed for the baseline survey. The total study population existed out of 11.980 households with at least one child aged under five years. After taking into account non-response and other factors (e.g. invalid data) in the calculations, the final desired sample size existed out of 409 households. For the purpose of this survey, 38 sub-districts were randomly selected from the six project districts. The sample size was distributed to each sub-district by the use of the population proportion method. In the end, baseline data was collected on 451 randomly selected households that live in villages in these 38 sub-districts.

The table below shows details on the distribution of the baseline respondents over the regions and districts that are targeted by the 3FM pneumonia project. It also shows the number of selected sub-districts per district.

Table 2. *Baseline survey - Number of respondents.*

Region	District	Nr. of sub-districts	Nr. of respondents
Amhara (*South Gondar zone*)	Simada	8	125 (27.7%)
	Ebenat	9	111 (24.6%)
	Total	***17***	***236 (52.3%)***
Oromia (*Guji zone*)	Girja	3	32 (7.1%)
	Gorodola	4	39 (8.6%)
	Total	***7***	***71 (15.7%)***
Somali *(Jijiga zone)*	Awbere	10	106 (23.5%)
	Gursum	4	39 (8.6%)
	Total	***14***	***145 (32.1%)***

Note: This table lists the baseline observations by district. The baseline survey is conducted by the ERCS with support from the NLRC. This survey was conducted in July 2017 to establish a baseline for the evaluation of the 3FM pneumonia project as a whole. Within the project districts (Column 2), 38 sub-districts were randomly selected (Column 3). In total, 452 households from these sub-districts have participated in the survey (Column 4).

Definition of the left-hand-sided variable

The baseline survey included a question on the health status of under-fives during the two weeks before the survey period. When children had been sick during the last two weeks, a follow-up question was asked on their specific symptoms or illnesses. Childhood pneumonia cases were defined as children who have suffered from fever and shortness of breath. The prevalence of pneumonia has been defined as a dummy variable, which is equal to one if an under-five in the household has suffered from pneumonia during the two weeks before the survey period. The variable equals zero otherwise. The prevalence rate of childhood pneumonia within the study sample is 19.7 percent.

Definition of the right-hand-sided variables

Since this analysis aims at researching which risk factors are significantly associated with the prevalence of pneumonia among the under-fives who are included in the baseline dataset, the established risk factors for childhood pneumonia will function as the right-hand-sided variables.

The risk factors that were found to be consistently associated with pneumonia among under-fives, presented in Section 2.2 (Table 1), have been selected as right-hand-sided variables. Furthermore, an additional potential risk factor in line with the aim of the 3FM Pneumonia project was selected as right-hand-sided variable; *Knowledge on pneumonia prevention*.

Subsequently, suitable indicators for these risk factors were identified from the baseline dataset (see appendix 1). For this end, I have used case definitions for multiple risk factors according to the systematic review of Jackson et al. (2013). Unfortunately, the baseline dataset does not include suitable indicators for all risk factors. I have only found suitable indicators for 15 out of the 25 identified risk factors. In addition, the indicators of *Health care costs*, *Maternal education* and *Child immunization coverage* are excluded from the regression model because they contain a substantial number of missing values.

Furthermore, two indicators are selected for the risk factor *Exposure to IAP* and the risk factors *Vaccine* pertussis and *Vaccine Hib* are represented by the same indicator (DPT-Hib-HepB vaccination).

The table below presents the descriptive statistics for the final set of variables that will be included in the analysis of the baseline dataset.

Table 3. *Baseline data - Descriptive statistics.*

	Variable	Description	Obs	Mean	Std. Dev.	Min	Max
Left-hand-sided variable	Prevalence of childhood pneumonia	Child has showed symptoms of pneumonia during last two weeks **[**0=No, 1=Yes**]**	451	0.197	0.398	0	1
Right-hand-sided variables	Breastfeeding practices	Child has been breastfed or not **[**0=No, 1=Yes**]**	452	0.841	0.366	0	1
	Symptoms malnutrition	Child has showed symptoms of malnutrition during the last two weeks **[**0=No, 1=Yes**]**	452	0.042	0.201	0	1
	Symptoms diarrhoea	Child has showed symptoms of diarrhoea during the last two weeks **[**0=No, 1=Yes**]**	452	0.092	0.291	0	1
	Symptoms measles	Child has showed symptoms of measles during the last two weeks **[**0=No, 1=Yes**]**	452	0.002	0.047	0	1
	Vaccine measles	Child has received measles vaccine **[**0=No, 1=Yes**]**	452	0.624	0.485	0	1
	Vaccine pertussis and Hib	Child has received DPT-Hib-HepB vaccine **[**0=No, 1=Yes**]**	452	0.710	0.454	0	1
	Maternal age	Age of mother in years	426	28.092	6.254	16	65
	Knowledge on pneumonia prevention	Number of pneumonia prevention measures mentioned by respondent	452	1.480	0.835	1	4
	Crowding level	Number of under-fives per room that is used for sleeping	452	1.498	0.700	1	5
	Indoor smoking	Regularity of smoking inside the house **[**0=Never, 1=Less than once a month, 2=Monthly, 3=Weekly, 4=Daily**]**	452	0.569	1.273	0	4
	Cooking location	Most common cooking location **[**0=Outside house, 1=Inside house**]**	452	0.274	0.447	0	1
	Biomass usage	Main cooking fuel used **[**0=Other than biomass, 1=Biomass**]**	452	0.743	0.437	0	1
	Local health care system	Perceived quality of care received at HF **[**0=Not so good, 1=Fair, 2=Good, 3=Very good**]**	452	1.892	0.821	0	3
	Access to health care	Time walking to HF **[**0= <60 minutes, 1= 60≤ X ≥120, 2= >120 minutes**]**	452	0.436	0.620	0	2
Control variables	Age of household head	Age of household head in years	452	35.763	8.677	19	65
	Education level of household head	Highest level of education completed by household head **[**1=Less than one year, 2=Primary, 3=Secondary, 4=Higher**]**	388	1.915	0.860	1	4
	Household headship	Specification of household headship **[**0=Couple-headed household 1=Single-headed household**]**	452	0.301	0.459	0	1
	Number of under-fives	Number of under-fives in the household	452	1.500	0.700	1	5
	Number of animals	Number of animals owned by household	452	5.383	7.571	-4	44

Note: This table describes the summary descriptives for all the variables that are included in the analysis of the baseline dataset. The table covers the left- and right-hand-sided variables as well as the control variables that are included in this analysis.
Source: ERCS & NLRC. (2017).

Empirical strategy

The baseline dataset will be analysed using multiple regression analysis. In order to assess the importance of the risk factors for childhood pneumonia, I will investigate the correlations between the above-mentioned indicators of these risk factors and the prevalence of childhood pneumonia.

I will do this by estimating the following logistic regression model:

$$Y_i = \beta_0 + \beta_1 Breastfeed_i + \beta_2 Malnutrition_i + \beta_3 Diarrhoea_i + \beta_4 MeaslesVaccine_i + \beta_5 HibVaccine_i + \beta_6 MaternalAge_i + \beta_7 KnowledgePneum_i + \beta_8 Crowding_i + \beta_9 IndoorSmoking_i + \beta_{10} Cooklocation_i + \beta_{11} Biomass_i + \beta_{12} QualityHF_i + \beta_{13} AccessHF_i + \beta_{14} SDFE_c + \beta_{15} C_i + \varepsilon_i.$$

In this regression model, Y_i stands for *Prevalence of childhood pneumonia.* While *Breastfeed, Malnutrition,…., AccessHF* stand for the selected risk factors of childhood pneumonia. The right-hand-sided variables as well as the left-hand-sided variable are measured at individual level *i* (*i = 1, ... , 391)*. Sub-district fixed effects $SDFE_c$ (c *= 1, ..., 30*) are included in the model to control for unobserved differences between sub-districts. Vector C_i contains a set of control variables, including household head's age and education level, household headship and the number of under-fives and animals. This regression model will be estimated with and without the sub-district fixed effects and control variables. To conclude, this regression model includes an idiosyncratic error term (ε_i).

3.2 Empirical results

Description of study population

This subsection describes the average respondents' characteristics, household and cooking practices. Descriptive statistics of the study population can be found in Appendix 2.

Most respondents are female and the vast majority of households covered by the baseline dataset is dual-headed. On average, households include 1.5 under-fives and the youngest child is 1.9 years old. In the average household, the father takes on the role of household head and the mother is the primary caregiver. The average household head is older (35.8 years vs. 28.5 years) and more educated than the average caregiver. Most primary caregivers have not completed more than one year of education, while most household heads have completed primary school.

When asked about preventive measures against pneumonia, a large majority of the households (85.62%) did not mention IAP reduction as a pneumonia prevention measure. Moreover, the majority of the participating households mainly use wood for cooking. From discussions with households from project villages, it can be concluded that the choice of cooking fuel strongly depends on the (natural) availability of cooking fuels. This is confirmed by the fact that the few survey respondents that have indicated to mainly use highly polluting animal dung for cooking, are all from the same zone.

Furthermore, the baseline data also indicates that cooking activities are usually not carried out in the house itself but in a separate building. In focus group discussions it became clear why it is not common to cook outside. Rain, strong winds and the presence of animals discourage outside cooking. In addition, especially during the evening, cooking outside is perceived to be dangerous in certain project areas. This could explain the substantial difference between single- and couple-headed households from the baseline dataset regarding this topic. Compared to couple-headed households, a substantial lower proportion of single-headed households have indicated to cook outdoors (5.9% vs. 17.1%).

Regression results

This section discusses the results of the three logistic regressions. The first regression model includes all right-hand-sided variables but no sub-district fixed effects and control variables. The second regression model is similar to the first but does also include sub-district fixed effects. The third regression model includes both sub-district fixed effects and a vector of control variables. For regression model (1) clustered standard errors (sub-district level) are estimated. Since regression models (2) and (3) include sub-district fixed effects, robust standard errors are estimated for these models.

For each regression model, I first discuss the estimation results of the model as a whole and then cover the results on the correlations between the right-hand-sided variable and the left-hand-sided variables separately. After estimation, all results were checked for multicollinearity using the Variance Inflation Factors (VIFs). This procedure indicated that no multicollinearity problems exist among the right-hand-sided variables.

Column (1) of Table 4 shows the standardized regression coefficients for the full model (sub-district fixed effects and control variables excluded). The model is found to be statistically significant at the 1% level (Chi-square = *95.96*). Nonetheless, the Pseudo-R^2 value is only 0.17. This means that 17% of the variation in the left-hand-sided variable is explained by this model. The regression results indicate that the following right-hand-sided variables are significant correlated with *Prevalence of childhood pneumonia* at the 1% level: *Measles vaccine* and *Knowledge on pneumonia prevention*. In addition, the correlation between *DPT-Hib-HepB vaccine* and *Prevalence of pneumonia* is significant at the 5% level.

Both the *DPT-Hib-HepB vaccine* (P<0.05) and the *Measles vaccine* (P<0.01) are strongly negatively correlated with *Prevalence of childhood pneumonia*. Their coefficients are quite large. The correlation coefficient of *DPT-Hib-HepB vaccine* takes a value of -0.38 standard deviations, while the correlation coefficient of *Measles vaccine* even takes the value of -0.53 standard deviations. The most obvious interpretation of these negative correlations would be that children have a lower chance of getting pneumonia if they have received (one of) these vaccines. This explanation would be in line with the literature since these vaccines have been found to play an important role in the protection of under-fives against pneumonia.

In addition, a significant positive correlation is found between *Knowledge on pneumonia prevention* and *Prevalence of childhood pneumonia* (P<0.01). The size of the coefficient is 0.64 standard deviations, which indicates a strong positive correlation. This coefficient suggests that children of caretakers who know more about pneumonia prevention measures, more often get pneumonia. In the first instance, this relationship sounds counterintuitive. The explanation for this relationship is not straightforward, since an endogeneity problem is in place here. One possible explanation would be that the chance of childhood pneumonia being recognised and diagnosed increases when knowledge on pneumonia improves. However, the underlying explanation can also be the other way around. Knowledge on pneumonia could also improve through information provision by health professionals when childhood pneumonia occurs. In that case, the occurrence of childhood pneumonia would positively affect the knowledge of caretakers on prevention measures. Unfortunately, it is impossible to be certain about the direction of the relationship between these two variables based on the available data.

The remaining right-hand-sided variables are not significantly associated with the *Prevalence of childhood pneumonia*. Their correlation coefficients are also relatively small.

Column (2) of Table 4 shows the standardized regression coefficients for the full model, including sub-district fixed effects. This model is also found to be statistically significant at the 1% level (Chi-square = *88.92*). In addition, the pseudo-R^2 value has doubled to 0.33. This means that the sub-district fixed effects alone explain about 16% of the variation in the left-hand-sided variable. This is comparable to the proportion of variation explained by all the other included variables together. Prevalence of childhood pneumonia thus seems to differ substantially between sub-districts. This can be explained, for example, by differences in the environmental contexts and the availability of health care services in the sub-districts.

Again, *DPT-Hib-HepB vaccine*, *Measles vaccine* and *Knowledge on pneumonia prevention* are found to be strongly correlated with *Prevalence of childhood pneumonia* (P<0.05). The (possible) processes in place are already discussed above. Furthermore, after including the sub-district fixed effects in the regression model, the correlation coefficients of *Malnutrition symptoms* and *Diarrhoea symptoms* became more significant. Their correlation coefficients are now relatively large at 0.36 (P<0.05) and 0.30 (P<0.10), respectively. These changes suggest a relatively strong association between these two variables and the sub-district factors (fixed effects). This would confirm the existence of strong regional differences in the health status of under-fives. The positive associations found between malnutrition and diarrhoea symptoms and the prevalence of childhood pneumonia is in line with the commonly found relationships between malnutrition, diarrhoea and pneumonia in children under five years of age.

Column (3) of Table 4 shows the standardized regression coefficients for the full model, including sub-district fixed effects and control variables. This model is also found to be statistically significant at the 1% level (Chi-square = 87.84). The pseudo-R^2 value has not changed compared to the second regression model.

The following control variables are included in this regression model: *Age of household head*, *Education of household head*, *Household headship*, *Number of under-fives* and *Number of animals*. After adding these control variables, the correlation coefficients of *Diarrhoea symptoms* and *Crowding* have become more significant and the correlation coefficient of *Measles vaccine* is no longer significant (P<0.05). These changes could be (partly) attributed to the fact that a relatively small number of observations is used for the estimation of this regression model.

Table 4. *Logistic regressions results.*

	Prevalence of childhood pneumonia		
	(1)	(2)	(3)
Breastfeeding [0=No, 1=Yes]	-0.039 (0.150)	-0.136 (0.210)	-0.105 (0.213)
Malnutrition symptoms [0=No, 1=Yes]	0.056 (0.122)	0.360** (0.147)	0.388** (0.160)
Diarrhoea symptoms [0=No, 1=Yes]	0.145 (0.175)	0.296* (0.178)	0.386** (0.184)
Measles vaccine [0=No, 1=Yes]	-0.532*** (0.149)	-0.489** (0.227)	-0.380 (0.277)
DPT-Hib-HepB vaccine [0=No, 1=Yes]	-0.377** (0.164)	-0.487** (0.221)	-0.559** (0.246)
Maternal age [In years]	-0.082 (0.192)	0.076 (0.216)	-0.171 (0.329)
Knowledge on pneumonia prevention [Nr. of measures mentioned]	0.639*** (0.213)	0.504** (0.207)	0.583** (0.241)
Crowding [Nr. of under-fives sleeping per room]	0.027 (0.140)	0.319 (0.194)	0.697** (0.337)
Indoor smoking [0=Never, 1=Less than once a month, 2=Monthly, 3=Weekly, 4=Daily]	0.144 (0.171)	0.020 (0.177)	-0.090 0.197
Cooking location [0= Outside house, 1=Inside house]	0.267 (0.189)	0.075 (0.185)	0.091 (0.202)
Biomass usage [0=Other than biomass, 1=Biomass]	-0.133 (0.196)	-0.042 (0.198)	-0.097 (0.213)
Perceived quality of health care [0=Not so good, 1= Fair, 2=Good, 3=Very good]	0.184 (0.172)	0.124 (0.206)	0.182 (0.242)
Distance to HF [0= <60 minutes, 1= 60≤ X ≥120, 2= >120 minutes]	-0.144 (0.214)	0.324 (0.223)	0.240 (0.245)
Sub-district fixed effects	No	Yes	Yes
Control variables	No	No	Yes
Constant	-1.760*** (0.213)	-1.839*** (0.486)	-2.086*** (0.577)
n	391	324	264
Chi-square (*df*)	95.96*** (13)	88.92*** (33)	87.84*** (37)
Psuedo R^2	0.171	0.325	0.329

Note: The left-hand-sided variable is the prevalence of pneumonia among under-fives. Childhood pneumonia cases were defined as children who have suffered from fever and shortness of breath.
The variable *Measles symptoms* was dropped since the left-hand-sided variable does not vary within the Measles symptoms = 1 category.
Control variables include age and education level of the household head, household headship and the number of under-fives and animals.
Standard errors in parentheses, with clustered standard errors (sub-district level) for Model (1) and robust standard errors for Models (2) and (3).
Significance is based on naive P-values.
** p < 0.10, ** p < 0.05, *** p < 0.01*

3.3 Discussion

The results indicate that the prevalence of childhood pneumonia differs substantially between project sub-districts. Overall, 19.7 percent of the under-five children from our baseline dataset have suffered from pneumonia during the two weeks before the survey period. This finding is comparable with the two weeks prevalence of pneumonia among under-five children (16.1%) found in the South Gondar zone by Fekadu, Terefe and Alemie (2014).

The regression results show that the prevalence of childhood pneumonia is strongly correlated with children's vaccination status. In line with the literature, the results indicate that children have a lower chance of getting pneumonia if they have received (one of) the following vaccines; DPT-Hib-HepB and measles vaccine. This emphasizes the need for increasing awareness regarding the importance of vaccinations in protection against childhood pneumonia. Therefore, it is important to discuss the topic of vaccination during the community conversation sessions that will be organised as part of the 3FM pneumonia project. The results also indicate that during these sessions, attention should be paid to the associations between malnutrition, diarrhoea and pneumonia in children under five years of age.

Additionally, the results show a positive correlation between knowledge on pneumonia prevention and pneumonia prevalence. This correlation suggests that children of caretakers who know more about pneumonia prevention measures, more often get pneumonia. It is possible that the likelihood of childhood pneumonia being recognised and diagnosed is strongly related to knowledge on pneumonia. Another explanation for this relationship would be that the occurrence of childhood pneumonia could lead to improved knowledge among caretakers on prevention measures through information provision by health professionals. Either way, this result emphasises the importance of project activities that focus on increasing knowledge and awareness on pneumonia signs and prevention measures. By the implementation of these activities, the RC aims to reverse the association between knowledge on pneumonia (prevention) and pneumonia prevalence.

Indoor air pollution reduction is one of the pneumonia prevention measures that will be promoted by the 3FM pneumonia project. The baseline data indicates that awareness on the role of IAP in the development of childhood pneumonia is generally very low. Only 14.4 percent of the respondents has mentioned IAP reduction as a pneumonia prevention measure. It is therefore vital that project communities will be informed on the danger of IAP and the ways in which they can reduce IAP. Since the use of a traditional stove as well as other traditional cooking practices are an important source of IAP, it is important to promote improved cooking practices through the 3FM pneumonia project.

Although the ICS intervention is mainly focused on ICS adoption and usage, the analysis of the baseline dataset indicates that other improved cooking practices should also be promoted alongside the introduction of improved stoves. Since the traditional cooking practices seem to vary slightly across the project districts, it is important to take these differences into account to ensure that suitable improved cooking practices can be promoted. For example, the choice of cooking location differs substantially across the districts. In addition, the choice of cooking fuel strongly depends on the (natural) availability of cooking fuels. The promotion of additional improved cooking practices should be adapted to the traditional cooking practices that are widespread in a certain area. Furthermore, the local traditional cooking practices and food preferences should be taken into account while designing and promoting ICSs.

3.4 Limitations

The main objective of this analysis is to answer RQ1: *Which risk factors are associated with the development of childhood pneumonia in the project areas?*. Because of multiple reasons, this analysis is of limited use for assessing the importance of the established risk factors for childhood pneumonia in the project areas.

First of all, the baseline dataset does not include suitable indicators for all established risk factors for childhood pneumonia. As a result, this limits the extent to which the importance of the risk factors in the development of childhood pneumonia in the study areas could be assessed.

Additionally, the extent to which the baseline survey accurately reflects the situation in the project areas can be questioned. For this survey, 38 sub-districts were randomly selected from the six project districts. Subsequently, 452 households from these 38 sub-districts were surveyed. Not all selected sub-districts are part of the project areas and not all project sub-districts were covered by the baseline survey. However, the project sub-districts that were not covered by the baseline survey are located in close proximity to the sub-districts that are covered by the baseline survey. Contextual factors are not expected to differ substantially across these sub-districts. Therefore, it was decided that the baseline data could be used to provide a rough indication of the baseline situation in the areas.

Furthermore, an important claim has to be made with regard to causality. Conducting causal research was not feasible given the available data. Hence, this analysis is limited to finding correlations. No causal claims can be made based on these correlations whatsoever. Consequently, serious endogeneity issues arise for various right-hand-sided variables of the regression model.

Moreover, data quality and internal validity could have been negatively affected by response bias that arose from the deployment of well-known RC volunteers as enumerators.

To conclude, the power of this analysis is threatened by the relatively small sample size of the baseline survey (n = 452). For eight sub-districts, less than ten observations were included in the dataset. The baseline survey thus has a very low coverage rate. The missing data problem, mainly caused by unanswered questions, reduces the sample size even further. The small number of observations used for this analysis results in a relatively low level of statistical power. This means that the analysis could have insufficient power to detect a significant correlation when there is one.

4. The 3FM pneumonia project

This section presents the results for RQ2: Which strategy is adopted by the RC to reduce pneumonia among under-fives in Ethiopia? To answer this research question, this section explores the strategy adopted by the RC to reduce childhood pneumonia in Ethiopia through the 3FM pneumonia project as well as the project activities that originate from this strategy. The section starts with a description of the selection process of the project areas. Subsequently, the components of the adopted strategy and the related project activities are described. The section concludes with introducing the areas and the stove model that are selected for the stove intervention.

4.1 Selection of project areas

Since the 3FM pneumonia project is being implemented by the ERCS with support from the NLRC, the project areas were selected in line with their partnership agreement. In this agreement the following priority regions were identified; Amhara, Oromia and Somali regions (See Figure 1). The project team has selected one target zone within each of these regions; South Gondar Zone (Amhara region), Jijiga Zone (Somali region) and Guji zone (Oromia region).

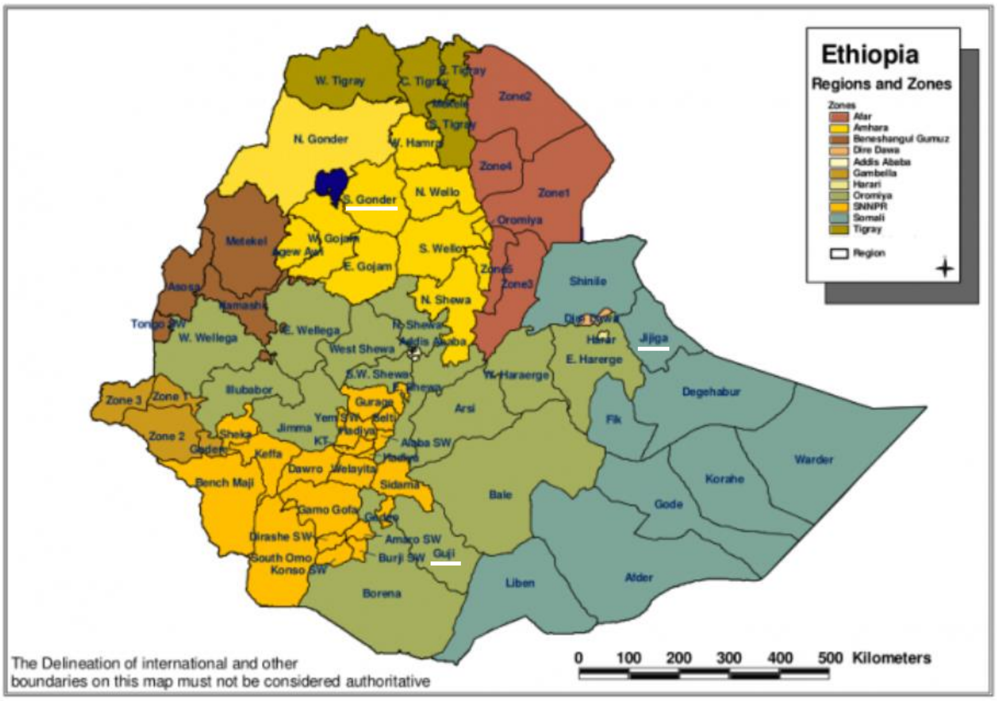

Figure 1. *The selected zones for the 3FM pneumonia project (World Food Programme & Disaster Prevention and Preparedness Commission, 2003)*

Subsequently, they selected two districts per zone based on the outcomes of an extensive vulnerability and resilience analysis. The capacity of the local RC branches and the local government were also taken into account in this process.

Using a phased implementation approach, this project will be introduced in approximately 46 rural sub-districts (Kebeles) spread over the six selected districts. Initially, the RC introduced the project in five sub-districts per selected district. The number of target sub-districts were scaled-up in the Amhara and Oromia regions during the second year of project implementation. These sub-districts were selected based on discussions with different sector ministries as well as the outcomes of the vulnerability and resilience analysis.

4.2 Project activities

The overall objective of the 3FM pneumonia project is to increase community resilience by reducing the mortality and morbidity rates of ARI among under-fives. The RC aims at reducing the negative effects of ARI on the daily lives and health of under-fives in Ethiopia by improving the capacity of health professionals and communities to diagnose, treat and prevent ARI.

The adopted strategy to reach the objective of this project is composed of three components. The first strategy component entails the strengthening of the case management of childhood ARI at health facilities. Related project activities exist out of the training of health facility staff, provision of essential medical equipment and introduction of a registration system. The second strategy component entails the improvement of the capacity of Health Extension Workers (HEWs) and RC volunteers to diagnose ARI and educate their communities on appropriate treatment of pneumonia. The project activities that are related to this strategy component include the training of these health professionals on various subjects, such as “Community Based Health and First Aid” (CBHFA) and pneumonia case management. The implementation of the project activities that are related to these two strategy components has started in 30 sub-districts in 2017. After approximately one year, when adequate pneumonia treatment was ensured in these sub-districts, the main emphasis shifted from treatment to prevention of pneumonia.

The third component of the adopted strategy focuses on increasing the capacity of communities to diagnose, treat and prevent ARI. An extensive Behavioural Change Communication (BCC) plan was designed in order to increase knowledge on risk factors for childhood pneumonia, to promote various preventive measures and to enhance health seeking behavior. Community conversation plays a central role in this BCC plan. These conversation sessions give the HEWs and RC volunteers a chance to share their newly gained knowledge on a large scale with the members of their communities. In the second quarter of 2018, they mobilized and invited the community for these conversation groups. From that moment on, they have been organizing and facilitating conversation sessions on a bi-weekly basis. Mainly women and influential persons such as community and religious leaders join these conversation sessions. All the selected discussion topics are directly or indirectly related to pneumonia. One of the conversation sessions covers IAP and ways in which it can be reduced. During this session, ICS usage is explicitly promoted.

In addition, the RC will implement a so-called ICS intervention. This was one of the few project activities that were not yet initiated at the moment that WUR was approached to conduct an impact assessment of the 3FM pneumonia project. Since various gaps exist in research on ICS adoption and impact of ICS usage, a study on this ICS intervention was expected to have considerable potential from an academic standpoint. The high popularity of ICSs among governments and donor agencies is another important reason for selecting the ICS intervention as a main focus for this impact study.

At the beginning of 2019, when the design of the impact study was finished, the ERCS initiated the ICS intervention. They organized training sessions for local women groups on ICS production and provided the women groups with building materials for the first rounds of production. The production and sale of ICSs will become an income-generating activity for the women of these local women groups. The ICSs will also be actively promoted by the ERCS, in line with the study design. More details on the experimental study design can be found in section 5.

4.3 ICS intervention – Areas of implementation

The ICS intervention was only introduced in the twenty sub-districts in South Gondar and Guji zone where the 3FM pneumonia project was initiated in January 2017. Details on these sub-districts can be found in Appendix 3. Due to security reasons, the ICS intervention was not introduced in the project sub-districts in Somali region.

Context

The Guji and South Gondar zone differ substantially. South Gondar zone is located 725 km away from Addis Ababa, in the sub-humid north-western part of the Ethiopian Highlands. Orthodox Ethiopian Christianity is the predominant religion in this zone. The sub-districts that are targeted by the ICS intervention in South Gondar zone are located in Simada and Ebenat district. These sub-districts are relatively large in terms of population (see Appendix 3). The road networks are generally very poor. Some of the target villages are not accessible during rainy seasons.

The semi-arid Guji zone, on the other hand, is located in South East Ethiopia. Negele, the zonal town, is located 570 km away from Addis Ababa. Protestantism is the predominant religion in this zone, followed by the Islam and Orthodox Christianity. The target sub-districts in Guji zone are located in the Gorodola and Girja districts. While the distance from Negele to the Gorodola district is approximately 30 km, the Girja district is located more than 200 km away from the zonal town. The target villages in these sub-districts are quite scattered and relatively small (see Appendix 3). Nearly all inhabitants of these villages are pastoralists. The distance from the target villages to the main road varies heavily and can get as high as 55 km. Most of these villages are only accessible by seasonal roads.

Prior NLRC projects

The NLRC has implemented multiple projects in the past on which the 3FM pneumonia project can build further, such as the Partners for Resilience (PfR) project (2009-2015), the 3FM neonatal project (2012-2016) and the 3FM WASH project (2014-2017). In two out of the five target sub-districts in Ebenat district, the PfR project as well as the 3FM neonatal project were implemented. In addition, most of the target sub-districts in Gorodola and Girja district (Guji zone) were targeted by the 3FM WASH project. Simada district, on the other hand, is a relatively new project area for NLRC. Its sub-districts were not targeted by any of the above-discussed projects.

4.4 ICS intervention – Stove

After a comprehensive assessment of suitable ICS models, the *Mirt* stove was selected for this ICS intervention (see Figure 2). The *Mirt* (translated as “best”) stove is highly specialized for *injera* baking, which is the main staple food in Ethiopia. This simple stove was developed by the Ethiopian Rural Energy Development and Promotion Center in the early 1990s (Yosef, 2007; Dawit, 2009). The Ethiopian government as well as various organisations, such as the Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ), have supported the production and commercialisation of the *Mirt* stove ever since (LaFave et al., 2018).

This stove can be produced from locally available raw materials and is relatively inexpensive (LaFave et al., 2018). Its market price ranges between 100 and 250 Birr (€3-8). Nevertheless, only 4% of all rural Ethiopian households owns this fuel-efficient cookstove (Gaia association, 2014).

Since the preparation of *injera* accounts for the majority of household fuelwood consumption (60%), the *Mirt* stove can lead to significant fuel savings for Ethiopian households (Bizzarri, 2010; Gebreegziabher et al., 2018). Various studies have been conducted to assess the fuel saving potential of the *Mirt* stove. While results of laboratory tests indicate fuel savings of approximately 50% (GIZ, 2011), field-based cooking tests have estimated the fuelwood savings potential of a *Mirt* stove to be 20% to 33% (Amare, Endeblhatu, & Muhabaw, 2015; Gebreegziabher et al., 2018). Estimations of fuelwood savings per kg of *injera* baked even range from 27% to 49.2% (Dresen et al., 2014; Workeneh, 2005; Yosef, 2007). According to Alemayehu et al. (2012), the *Mirt* stove saves 290 kg of fuelwood annually per household member compared to the traditional stove.

However, the fuel savings associated with the *Mirt* stove are not self-evident. Increasing efficiency gains over time indicate that people need to adapt to the new cook technology (Bluffstone et al., 2015).

The actual impact of *Mirt* stove usage on Indoor Air Quality (IAQ) and respiratory health has remained relatively unexplored. Although results of laboratory tests suggest that the *Mirt* stove can reduce emissions by 50 percent compared to the traditional stove (GIZ, 2013), field studies on this topic are very scarce. In addition, only a few studies have assessed the health effects of the *Mirt* stove. These studies did not find significant improvements in respiratory health of children (LaFave et al., 2018). All in all, it is not yet clear if the health benefits of the *Mirt* stove are relatively small or if its health improving potential is attenuated through, for example, continued use of traditional stoves alongside the *Mirt* stove. This should not be necessary since *injera* baking and other cooking and boiling activities can be performed simultaneously on the *Mirt* stove (Dresen et al., 2014). However, the fact that *injera* is in general not baked on a daily basis could explain why households keep using traditional stoves besides the *Mirt* stove.

The RC is exploring the possibilities for integrating a chimney in the design of the *Mirt* stove to ensure that the use of the stove would result in substantial improvements in IAQ and child health. An approved design of a suitable add-on chimney exists but there is no evidence when it comes to the risk of fire. Therefore, the ERCS first has to establish evidence on the incidence of fire while using the *Mirt* stove with this add-on chimney before the chimney can be produced and distributed.

Figure 2. *The Mirt stove design (Forests News, 2014)*

5. Study design of strategies to increase ICS adoption

This section discusses the experimental design of the impact study on the ICS intervention in more detail. The objective of this study is to answer RQ3: How effective are stove subsidies and information provision on stove benefits in increasing ICS adoption?. This section first describes the treatments of the RCT. Followed by an introduction of the selected impact indicators and hypotheses. Next, the randomization process and the research population of the trial are described. Subsequently, the sample size, statistical power, empirical strategy and data collection process of the study are covered. To conclude, the ethical considerations, limitations and planning of the study are reviewed.

5.1 Treatments

The RCT will be conducted as an integrated part of the ICS intervention. The ERCS has already started with the training of local women groups in *Mirt* stove production. When they have produced a specific number of stoves, the ERCS will organise stove introduction meetings. Per sub-district, two meetings will be organised. The objective of these meetings is to inform households on the *Mirt* stove. At these meetings, attendees will also be given the opportunity to buy the stove directly. Afterwards, women groups will continue the production and sale of *Mirt* stoves.

The treatment factors of this RCT relate to the design of these stove introduction meetings. We will vary the information provision on the benefits of *Mirt* stove usage as well as the access of participating households to free stoves. Our primary aim is to study how these factors affect ICS adoption.

<u>Information provision</u>

As discussed earlier, the mismatch between stove (choice) preference and intra-household bargaining power form a barrier to ICS adoption. Women generally tend to have stronger preference for ICSs than men, but they usually lack the authority to make the actual stove purchase decision (Miller & Mobarak, 2013). Therefore, it is crucial to convince men of the benefits of ICS adoption in order to motivate them to support their wives in this adoption decision. Usually the RC would mainly focus on the health effects of the *Mirt* stove. However, men are expected to value economic benefits higher than health benefits. This is due to the fact that health benefits of the *Mirt* stove are mainly experienced by women and children while economic benefits are directly experienced by men.

We will vary the provision of information on *Mirt* stove benefits between the two stove introduction meetings that will be organised in each sub-district. This enables us to assess how information provision on *Mirt* stove benefits affects stove adoption through intra-household bargaining processes. At one meeting, we will primarily focus on the expected health benefits of the *Mirt* stove. At the other meeting, we will discuss both the expected health and economic benefits of the *Mirt* stove.

<u>Free distribution of Mirt stoves</u>

Since most households are credit constrained and chronically short on cash, the decision to adopt an ICS is highly responsive to stove price (Bensch et al., 2015; Lewis & Pattanayak, 2012; Mobarak et al., 2012). In reaction to households' financial barriers, vouchers for free *Mirt* stoves will be handed out at the end of the stove introduction meetings. These vouchers represent a 100% stove subsidy and can immediately be exchanged for a free *Mirt* stove. However, not all households will receive this voucher.

2x2 factorial design

With these two treatment factors, four treatment conditions can be created (Table 5). This 2x2 factorial design enables us to assess the main effects of the individual treatment factors as well as potential interaction effects between these factors.

Table 5. *The 2x2 factorial study design*.

	No free stove (No subsidy)	**Free stove** (100% subsidy)
Stove introduction meeting focused on health benefits	Treatment 1	Treatment 2
Stove introduction meeting focused on health and economic benefits	Treatment 3	Treatment 4

Note: This table gives an overview of the two treatment factors and four treatment conditions that are part of the 2x2 factorial design. The treatment factors include the free distribution of ICSs and information provision on the benefits of the *Mirt* stove.

Stove subsidies

To learn more about the relationship between the level of stove subsidy and ICS adoption, we will introduce multiple subsidy levels in this RCT. Based on the literature, we expect a positive linear relationship between the level of stove subsidy and the saturation level of ICSs. With full-cost pricing resulting in extremely low adoption rates and handing out free stoves resulting in relatively high adoption rates. However, we want to assess if it is necessary to give a 100% stove subsidy in order to increase the rate of adoption substantially. If a lower subsidy would already be effective in increasing stove adoption, this would be a more cost-effective strategy to increase ICS adoption rates. To learn more about the relationship between the subsidy level and stove adoption, vouchers for different levels of stove subsidies will be issued. Four different vouchers will be issued, representing 0%, 40%, 70% and 100% subsidies. The results regarding this treatment factor will inform us on the optimal level of subsidy, taking into account the adoption rates and sustainability of stove usage at the different subsidy levels. This 2x4 factorial design enables us to assess the effects of the individual treatment factors, information provision and subsidy level as well as the potential interaction effect between these factors (See table 6).

Table 6. *The 2x4 factorial study design.*

	No subsidy	**40% subsidy**	**70% subsidy**	**100% subsidy**
Stove introduction meeting focused on health benefits	Treatment 1	Treatment 2	Treatment 3	Treatment 4
Stove introduction meeting focused on health and economic benefits	Treatment 5	Treatment 6	Treatment 7	Treatment 8

Note: This table specifies the extended study design; the 2x4 factorial design. By introducing two additional levels of subsidy, the treatment factor on the free distribution of ICSs (from the 2x2 factorial design) is changed to a treatment factor on stove subsidy. This also leads to an increase from four to eight treatment conditions..

Treatment groups 1 to 4 form the control group for the treatment factor on information provision. For the treatment factor on the free distribution of Mirt stove, treatment groups 1 to 3 and 5 to 7 form the control group. The control group for the treatment factor on stove subsidies is formed by treatment groups 1 and 5, since households from villages that are assigned to these treatment groups will receive no stove subsidy at all.

5.2 Impact indicators

This sub-section discusses the impact indicators that are selected for this study. First the main impact indicator is defined and then a set of secondary impact indicators is specified. For each impact indicator it is mentioned which survey question(s) will be asked to collect data on the specific indicator, except for IAQ and exposure to IAP. Data on these indicators will be collected by the use of measurement devices.

ICS adoption

The main impact indicator of this study is the rate of ICS adoption, which is calculated as the share of households that owns a *Mirt* stove. The household member responsible for cooking will be asked to specify the types of stoves owned by the household.

Stove utilisation rate

The utilisation rate of ICSs is indicated by the proportion of individual dishes prepared with the *Mirt* stove. To account for the possibility that multiple stoves may be used (simultaneously) to prepare one single meal, we define an individual dish as one component of a meal that is prepared on a stove. Enumerators will crosscheck ICS usage as part of the interview by verifying if the *Mirt* stove has recently been used.

Firewood consumption

To record firewood consumption, the household member responsible for cooking will be asked to specify the meals cooked on a typical day and the types of stoves generally used for the preparation of these meals. For each stove application, we will record the cooking fuel type used. In case of firewood usage, the respondent will be asked to pile up the amount of firewood used on a typical day. The resulting wood pile will be weighed with scales. In combination with information on variation in firewood consumption throughout a typical week, this data will be used to determine the weekly household consumption of firewood.

Time allocation patterns

To record the amount of time spent on cooking and firewood collection within a household, we will ask the household member responsible for cooking to indicate how much time each of the household members spent on these activities. Subsequently, we will aggregate the time spent by individual household members on these activities. One of the resulting indicators will indicate the self-reported amount of time spend by households on cooking on a typical day. The other resulting indicator will specify the total amount of time spend by households on firewood collection in the course of a week.

Indoor air quality

Indoor air quality will be indicated by the levels of indoor air pollutants. We will measure the static CO and $PM_{2.5}$ concentrations in kitchens by the use of Particulate Matter (PM) sensors and CO data loggers.

Exposure to indoor air pollution

The exposure of under-fives to (indoor) air pollution is indicated by the quality of air that they are exposed to on a daily basis. We will attach passive diffusion tubes to the clothes of children to measure the levels of indoor air pollutants that they are exposed to during a 24-hour time period.

Child health

We will research the impact on child health by assessing the prevalence of diseases that are strongly related to IAP exposure. For this purpose, we focus on symptoms of respiratory diseases and eye problems. The caretaker of the household will be asked if any of the children under five years of age has suffered from symptoms of respiratory diseases or eye problems during the two weeks before the survey period. For respiratory diseases, we will check for the following symptoms; cough, asthma, fever and difficulty in breathing. The prevalence of respiratory diseases and eye problems among under-fives will thus be elicited on a self-reporting basis. The two resulting health-related variables will specify the number of under-fives per household that have suffered from symptoms of respiratory diseases and eye problems during the two weeks before the survey period.

Furthermore, we will record the pneumonia-specific mortality rate in under-fives. The caretaker of the household will be asked if there were deaths among under-fives in their household during the last year that can be attributed to pneumonia. We will calculate the pneumonia-specific mortality rate in under-fives by dividing the total number of pneumonia deaths in under-fives from the surveyed households by the total number of under-fives from these households during the same time frame. We will multiply the outcome by 1000 to express the mortality rate in number of deaths per 1000 children per year.

5.3 Hypotheses

The main aim of this study is to assess if the previously discussed treatment factors (See Sub-section 5.1) truly increase the rates of ICS adoption. We will use one-tailed hypothesis tests since there is no reason to believe that the treatment factors would lead to a reduction in ICS adoption rates. We will test whether access to free stoves and information provision on economic benefits of the *Mirt* stove significantly increase ICS adoption. We will also test for an interaction effect between the free distribution of ICSs and information provision. In addition, we will research the relationship between subsidy level and ICS adoption.

The following four hypotheses will be tested:

H1: Information provision on the economic benefits of ICSs increases stove adoption.
(ICS adoption rates: Treatment groups 5-8 > Treatment groups 1-4)

H2: Free distribution of ICSs leads to increased stove adoption.
(ICS adoption rates: Treatment groups 4 & 8 > Treatment groups 1-3 & 5-7)

H3: Subsidy levels increase ICS adoption.
(ICS adoption rates: Treatment groups 4 & 8 > Treatment groups 3 & 7 > Treatment groups 2 & 6 > Treatment groups 1 & 5)

H4: A positive interaction effect exists between free distribution and information provision.
(ICS adoption rates: Treatment group 8 minus Treatment groups 5-7 > Treatment group 4 minus Treatment groups 1-3)

These four hypotheses will also be tested regarding the secondary outcome variables of this study; *Mirt* stove utilisation rate, fuel consumption, time spent on firewood collection, IAP (exposure), prevalence of childhood pneumonia and eye problems and pneumonia-specific mortality among under-fives.

5.4 Randomization

We chose to use a cluster randomised design to evaluate the impact of the discussed treatment factors on ICS adoption. Treatment will be assigned by use of a block randomisation procedure. This means that first clusters (villages) will be stratified by sub-district and distance to central point in sub-district. Subsequently, the treatment conditions will be randomized on village level. This randomization process will result in eight groups of villages that are expected to be balanced in the observed and unobserved characteristics that could influence the treatment effects. Furthermore, on average, balance will be ensured in the baseline characteristics of the target population in these eight groups of villages. By means of this randomization process we can thus eliminate bias in treatment assignment.

5.5 Research population

The ICS intervention will target approximately 483 villages in 20 sub-districts. The treatment conditions of this RCT will be randomly assigned among these villages. Since the goal of the 3FM pneumonia project is to reduce pneumonia-related childhood mortality, the target population of the ICS intervention consists out of households with at least one child under-five years of age. In total, 10.736 households with at least one child under five years of age live in the 483 study villages. However, on request of the RC, the 10 percent most vulnerable households with under-fives will be excluded from trial participation. These households will receive a free *Mirt* stove regardless of the treatment condition that will be assigned to their village. As a result, a total of 9.662 households with under-fives will be included in the trial. The research population thus consists out of this subset of households.

5.6 Sample size and statistical power

It is essential to assess the statistical power of an impact assessment before the assessment is conducted. The statistical power of a study indicates its ability to detect differences in the outcome variables, if these differences would exist in reality. Fundamental ingredients of power calculations are the desired level of power, significance level, baseline mean value (and standard deviations), ICC and either sample size or the expected treatment effect. For the main outcome variable (ICS adoption), we have set the critical significance level (α) at 0.05, by convention. For the secondary outcome variables, we have adjusted for multiple testing by using a Sidak-adjusted alpha. Taking into account the number of hypotheses tested, the critical significance level (α) is set at 0.0064 ($\alpha_{SID} = 1 - (1 - 0.05)^{1/8}$). We have set the level of power at 0.80, as is convention. This level of power implies a chance of 80% to find a statistically significant difference between groups in the research population, if it does exist.

With the level of power specified, you can either use the expected effect to find the necessary number of respondents to detect this effect or you can use the number of respondents to research the Minimal Detectable Effect (MDE). This sub-section aims at exploring the MDE sizes for the selected outcome variables of this study. The MDE expresses the smallest average treatment effect that is detectable for the given level of power and significance.

We have used the 2x2 factorial study design for these power calculations, with the free distribution of ICSs and information provision on stove benefits as treatment factors. We have calculated the minimal detectable main effects of the treatment factors as well as the minimal detectable interaction effects between these treatment factors. Table 7 shows the results of these power calculations grouped by outcome variable. Panel A includes the main outcome variable and the other panels include the secondary outcome variables. For most of the secondary outcome variables, results of multiple studies have been included in the table.

The outcome variables of the study are divided in five categories. In the first column, the outcome variables are specified. In the second column, we refer to the authors of prior studies that have assessed the impact of comparable ICS interventions on these specific outcome variables. The mean value of the specific outcome variable in their control group are mentioned in the third column. The treatment effects that they have found on these outcome variables are specified in the fourth column. Next, the baseline data collected by the RC on the outcome variables is explored. Unfortunately, only two of the selected outcome variables are covered by the baseline dataset. The mean value and Intraclass Correlation Coefficient (ICC) of these two variables from the baseline dataset are reported in the fifth and sixth columns. The ICC describes how strongly households in the same cluster (village) resemble each other. Based on the ICCs of the variables from the baseline dataset, the ICCs of the outcome variables that were not included in the baseline dataset are set at 0.2 for these power calculations. Furthermore, the reported mean values in the control groups of prior studies are taken as baseline mean values for the outcome variables that were not included in the baseline dataset.

Since this study design includes cluster randomization, the power calculations need to take into account the number of clusters used. Based on the results of the power calculations, we decided that the endline survey would cover only 264 of the 483 clusters (villages). As half of these 264 clusters will be assigned to the treatment conditions that are not part of the 2x2 factorial design, a total of 132 clusters are used for these power calculations. Furthermore, these calculations are based on 3 respondents per cluster. The power of the study would not increase substantially when we would increase the number of clusters or respondents per cluster.

The results of the power calculations are shown in the last two columns of Table 7. The minimal detectable main effects of the treatment factors are specified in the seventh column. We have compared these MDEs to the effects found by prior studies. We will not be able to detect the expected main effects of the treatment factors on a particular outcome variable if the MDE is larger than the expected effect. At least, not with the specified levels of power and significance and numbers of clusters and respondents. We have underlined the MDEs that are smaller than the expected main effects of the treatment factors. Based on the results of these power calculations, we expect to be able to detect these treatment effects. Nevertheless, we have to keep in mind that the actual impact of the treatment factors can deviate from our expectations.

Furthermore, we have assessed the minimal detectable conditional marginal effects of the two treatment factors. Conditional marginal effects consist out of additional interaction effects between the two specified treatment factors. We consider the existence of interaction effects between the treatment factors to be quite plausible. These effects are harder to detect as only one out of the four groups of clusters will receive a treatment existing out of both treatment factors, while the individual treatment factors will be both provided to two out of the four groups of clusters.

The results of the power calculations indicate that we will be able to detect a main effect of 10 percentage points or more on the main outcome variable (Panel A), with 33 clusters per treatment arm. Since the RC project team expects the main effects of the treatment factors on ICS adoption to be approximately 20 percentage points, it will definitely not impossible to detect the expected main effects on ICS adoption. In order to detect a potential interaction effect, the mean in the treatment group has to increase by at least another 16 percentage points.

Furthermore, the results of the power calculations indicate that we will be able to detect the expected main effects of the treatment factors on most of the outcome variables from Panel C, by means of this study design. In order to assess the impact of the treatment factors, through ICS stove adoption, on the outcome variables from Panel D and E, we need to conduct relatively expensive IAQ measurements. Based on the outcomes of prior studies, we conclude that is slightly harder to measure impact on indoor $PM_{2.5}$ concentrations compared to indoor CO concentrations. Most main effects of treatment factors on IAQ that can be detected with measurements in 3 households per cluster, can also be detected with measurements in 'only' 1 household per cluster (indicated by two asterisks in Table 7). These results indicate that we should consider random selection of one of the three surveyed households per cluster for IAQ measurements.

In line with our expectations, the power calculations indicate that it will be relatively hard to detect the impact of the treatment factors on child health (Panel F). While we might be able to detect the main effects of the specified treatment factors on childhood pneumonia, this will most certainly not be the case for pneumonia-specific mortality among under-fives. This is due to the fact that the pneumonia-specific mortality rate among under-fives is very low and the expected treatment effects are relatively small. As a result, we can only detect a reduction of 3 percentage points or more while the mean baseline pneumonia-specific mortality rate in the literature is no more than 1.75 percent. This implies that we would solely be able to detect a negative treatment effect on pneumonia-specific mortality if a reduction of more than 100 percent could be realised. Since this is practically impossible, we expect to be not able to detect any treatment effects on pneumonia-specific mortality among under-fives.

Table 7. *Power calculations.*

Outcome measures	Source	Reported mean value in control in source (SD)	Reported difference (treatment effect)	Mean value in RC baseline (SD)	ICC	Main effects of treatment factors [N_j=132 (2*66), N_i=3]	Interaction effects between treatment factors [N_j=132 (4*33), N_i=3]
PANEL A: Stove adoption						**Minimal detectable effects**	
Adoption of ICSs	Expectations of RC expectation	6%	+20 pp	-	0.20*	+10 pp	+16 pp
PANEL B: Stove usage							
Utilisation rate of ICSs	Bensch & Peters, (2015) n= 250	0.7%	+68.4 pp	-	0.20*	+9 pp	+17 pp
PANEL C: Household cooking efficiency							
Time spent on cooking (per day in minutes)	Bensch & Peters, (2015) n= 229	333	-82 (21.8)	-	0.20*	-29.41 minutes	-49.97 minutes
Firewood consumption (per week/household in kg)	Bensch & Peters, (2015) n= 228	87.58 (70.67)	-26.78	-	0.20*	-30.22 kg	-43.06 kg
Firewood consumption for *injera* baking (per year/capita in kg)	Dresen et al., (2014) n= 148	399.8 (195.2)	-155.3	-	0.20*	-83.46 kg	-118.95 kg
Firewood consumption for non-*injera* cooking (per y/c in kg)	Dresen et al., (2014) n= 148	657.3 (314.7)	-58.8	-	0.20*	-134.56 kg	-191.77 kg
PANEL D: Indoor air quality							
Concentration of $PM_{2.5}$ (in mg/m3)	Dutta, Shields, Edwards, & Smith, (2007) n= 87	1.25 (1.61)	-0.31	-	0.20*	-0.69 mg/m3	-0.98 mg/m3
Concentration of $PM_{2.5}$ (in mg/m3) [48-h average]	Masera et al., (2007) n=33	1.02 (0.79)	-0.68 (0.27)	-	0.20*	-0.49 mg/m3**	-0.48 mg/m3
Concentration of $PM_{2.5}$ (mg/m3) [48-hr measurements]	Chengappa et al., (2007) n=15	0.65 (1.01)	-0.29	-	0.20*	-0.43 mg/m3	-0.62 mg/m3
Concentration of CO (in ppm)	Dutta, Shields, Edwards, & Smith, (2007) n= 98	10.82 (8.71)	-4.17	-	0.20*	-3.72 ppm	-5.31 ppm
Concentration of CO (in ppm) [48-h average]	Masera et al., (2007) n=32	8.88 (4.44)	-5.86 (2.66)	-	0.20*	-1.90 ppm**	-2.71 ppm
Concentration of CO (in ppm) [48-hr measurements]	Chengappa et al., (2007) n=15	8.67 (7.80)	-5.99 (2.8)	-	0.20*	-3.34 ppm**	-4.75 ppm
PANEL E: Exposure to indoor air pollution							
CO Child (ppm hours) [24-hr measurements]	Barnes, Mathee, & Thomas, (2011) n=74	3.46 (2.70)	-1.61 ppm hours (2.50)	-	0.20*	-1.15 ppm hours	-1.65 ppm hours
PANEL F: Child health							
Coughing on a daily basis	Pine et al., (2011) n= 259	19.2%	-16.2 pp	21.7% (0.41)	0.24	-15 pp	-19 pp
Self-reported ARI signs among under-fives (during last two weeks)	-	-	-	19.6% (0.40)	0.22	-14 pp	-18 pp
Pneumonia-specific mortality among under-fives	Bang et al., (1990) N= 43, n= 10.123	1.75%	-0.94 pp	-	0.20*	-3 pp	-3 pp

Note: This table gives an overview of the data used for the power calculations as well as the outcomes of these calculations. For these calculations, the 2x2 factorial study design is used, with the free distribution of ICSs and information provision on stove benefits as treatment factors. These calculations are based on 132 clusters and 3 observations per cluster.

The first column specifies the outcome variables and the second column refers to the authors of prior stove studies that have focused on these variables. The mean values in the control groups and found treatment effects by these studies are specified in the third and fourth column. The fifth and sixth columns contain the mean values and ICCs of the outcome variables from the RC baseline dataset. The seventh column specifies the minimal detectable main effects of the treatment factors. The MDEs that are smaller than the expected treatment effects are underlined. The last column specifies the minimal detectable interaction effects between the two treatment factors.

For panel A : $\alpha = 0.05$; For the other panels : $\alpha_{SID} = 0.0064$. Power = 0.8, nr. of treatment arms = 4.
*= These ICCs are assumed to be 0.2; **= These MDEs can be detected with measurements in 1 household instead of 3 households per cluster

Overall, these power calculations (based on the 2x2 factorial design) provide a very optimistic estimation regarding the MDEs on the secondary outcome variables. This is due to the fact that the adoption rate of the *Mirt* stove is expected to be relatively low among households that do not receive a free stove. Therefore, we expect the power of this study to detect treatment effects on the secondary outcome variables, through ICS stove adoption, to be relatively low. By the use of a larger sample, it would be possible to detect slightly smaller effects on the secondary outcomes.

5.7 Empirical strategy

This sub-section discusses the strategies adopted to identify treatment effects and test the previously formulated hypotheses (See Sub-section 5.2). In addition, the variables are specified on which we will focus when we test for heterogeneous treatment effects.

Identification strategy

We will employ two approaches to estimate the impact of the treatments of this RCT and to test the formulated hypotheses. We will begin by estimating the intent-to-treat (ITT) effects. The ITT effect is simply the difference in mean values of the outcome variables for the treatment and control groups. In the case of this study, the ITT effects represent the effects of inviting households to one of the two stove introduction meetings and offering them one of the four subsidy vouchers.

However, we believe that there will be subjects who fail to comply with the assigned treatment. This phenomenon is called non-compliance. In the case of this study, compliance is defined as being present at the designated stove introduction meeting and receiving the designated stove voucher. We believe that there will be a considerable share of households that will not be present at the stove introduction meeting for which they will be invited. In this case, households do not receive information on the *Mirt* stove nor a subsidy voucher. Take-up in the control group is very unlikely since there is very restricted access to the meetings and vouchers in the control group. With a share of the households expected to fail to receive the voucher or visit the meeting that is part of the treatment that has been assigned to them and no expected uptake of vouchers or presence at meetings that are not part of the assigned treatment, we expect one-sided non-compliance. In order to correct for the non-compliance in treatment groups, we will calculate the effect of Treatment-on-the-Treated (TOT). The TOT effects can be estimated by dividing the ITT effects by the compliance rate. The TOT effect thus represent the average treatment effect on compliers and thereby the actual treatment effects. We will estimate compliance rates based on the data from the attendance sheets of the stove introduction meetings. These documents will contain details on the presence of, and voucher receipt by, the research population.

We will assess the ITT effects of the treatment factors on the main outcome variable (ICS adoption) and test the related hypotheses (See Sub-section 5.2) by use of the linear probability models presented below. With the binary outcome variable (Y_{iv}) indicating ownership of the *Mirt* stove on household level in village v ($v = 1, ..., 483$). This variable will be equal to one if household i owns a *Mirt* stove during the survey period. The variable equals zero otherwise. In these regression models, T is a dummy variable on treatment assignment and its parameter represents the estimated ITT effect of the specified treatment factor on ICS adoption. Since the randomization is stratified on village level, the regression models include village fixed effects (α_v). To conclude, the models also include the usual idiosyncratic error term (ε_i).

First, we will test whether information provision on the economic benefits of the *Mirt* stove increases ICS adoption (Hypothesis 1). In order to test this, we will compare the ICS adoption rates in groups that were invited for a stove introduction meeting at which economic benefits were discussed (treatment groups 5 to 8) with the adoption rates in groups that were invited to a meeting at which these economic benefits were not discussed (treatment groups 1 to 4).

We will compare the adoption rates in these different treatment groups by means of the following regression model:

$$Y_{iv} = \beta_0 + \beta_1 T^{5\text{-}8}{}_v + \alpha_v + \varepsilon_i \qquad (1)$$

In which the dummy variable $T^{5\text{-}8}$ will equal one if a household has been assigned to treatment group 5, 6, 7 or 8. The variable equals zero otherwise. The parameter of interest is β_1. This parameter is hypothesised to be positive. We will run this regression with the full sample.

Next, we will test whether the free distribution of ICSs increases ICS adoption (Hypothesis 2). We will do so by comparing the ICS adoption rates in groups that were offered a free stove (treatment groups 4 & 8) with the ICS adoption rates in groups that were not offered a free stove (the remaining treatment groups) by means of the following regression model:

$$Y_{iv} = \delta_0 + \delta_1 T^{4,8}{}_v + \alpha_v + \varepsilon_i \qquad (2)$$

In which the dummy variable $T^{4,8}$ will equal one if a household has been assigned to treatment group 4 or 8. The variable equals zero otherwise. The parameter of interest (δ_1) is hypothesised to be positive. Again, we will run this regression with the full sample.

We hypothesise subsidy levels to increase ICS adoption (Hypothesis 3). In order to learn more about this relationship, we will compare the ICS adoption rates across the subsidy groups. By means of the regression model below, we will compare ICS adoption rates across groups that were offered a 0% subsidy voucher (treatment groups 1 & 4), 40% subsidy voucher (treatment groups 2 & 6), 70% subsidy voucher (treatment groups 3 & 7) and 100% subsidy voucher (treatment groups 4 & 8).

$$Y_{iv} = \gamma_0 + \gamma_1 T^{2,6}{}_v + \gamma_2 T^{3,7}{}_v + \gamma_3{}^{4,8}{}_v + \alpha_v + \varepsilon_i \qquad (3)$$

This regression model includes three dummy variables on treatment assignment. If a household has been assigned to treatment group 2 or 6, dummy variable $T^{2,6}$ equals one and the other dummy variables equal zero. The same applies if a household has been assigned to treatment group 3 or 7. In that case, the dummy variable $T^{3,7}$, equals one and the other dummy variables equal zero. If a household has been assigned to treatment group 4 or 8, dummy variable $T^{4,8}$ equals one and the other dummy variables equal zero. Since treatment groups 1 and 5 are used as reference category, all three dummy variables equal zero if a household has been assigned to treatment group 1 or 5. The parameters of interest are γ_1, γ_2 and γ_3. These parameters are hypothesised to be positive and increasing in size ($\gamma_3 > \gamma_2 > \gamma_1$). We will run this regression with the full sample.

To conclude, we will test the hypothesised positive interaction effect between free distribution of ICSs and information provision (Hypothesis 4). In order to so, we will first calculate the differences between the ICS adoption rates in the groups that were offered a free stove with the groups that were not offered a free stove but were invited for the same stove introduction meeting (treatment group 4 minus treatment groups 1, 2 & 3 and treatment group 8 minus treatment groups 5, 6 & 7). Next, we will compare the resulting differences. We will use the regression model below to calculate these differences in ICS adoption rates. This regression model includes three dummy variables on treatment assignment. Since treatment groups 1 to 3 are used as reference category, the constant term will capture the effect of assignment to these treatment groups.

$$Y_{iv} = \zeta_0 + \zeta_1 T^4{}_v + \zeta_2 T^{5\text{-}7}{}_v + \zeta_3 T^8{}_v + \alpha_v + \varepsilon_i \qquad (4)$$

The parameters of interest are ζ_0, ζ_1, ζ_2 and ζ_3. The difference between ζ_3 and ζ_2 is hypothesised to be larger than the difference between ζ_1 and ζ_0. We will run this regression with the full sample.

To control for non-compliance, we will also report the TOT effects of the treatment factors on ICS adoption. The TOT parameter is equal to the ITT parameter divided by the estimated compliance rate. This parameter thus represents the average magnitude of the impact of the treatment factors on compliers.

Furthermore, we will also use the four regression models specified above to study the impact of the treatment factors on the secondary outcome variables. The secondary outcome variables of this study are *Mirt* stove utilisation rate, fuel consumption, time spent on firewood collection, IAP (exposure), prevalence of childhood pneumonia and eye problems as well as pneumonia-specific mortality among under-fives.

Heterogeneous treatment effects

When testing for heterogeneous treatment effects, we focus on the exogenous variables mentioned below. We expect these variables to affect the effects of the treatments without being affected by the treatments themselves.

We expect to find a positive association between the following variables and the ICS adoption rate:

- Age of caregiver. The experience of caregivers and their awareness on the importance of ICS adoption is expected to increase with age. Therefore, we expect the ICS adoption rate to be higher among households of which the caregiver is older compared to households with very young caregivers.
- Household income. Higher-income households are more likely to be able to afford stove. They are also expected to spend their money more easily than lower-income households. As a result, we expect the ICS adoption rate to be higher among households that have a relatively high income, compared to lower-income households.
- Education of household head and caretaker. More educated households heads and caretakers are expected to have stronger understanding of the dangers of IAP and the necessity of adopting an ICS. They are expected to be better able to make an informed choice on ICS adoption. Education status might also positively influence the ability to accept new ideas and having higher confidence in dealing with health professionals (Tugumisirize, Tumwine, & Mworoza, 2002). Consequently, we expect the ICS adoption rate to be higher among households with more educated household heads and caretakers, compared to households with less educated households heads and caretakers.
- Distance to closest firewood collection site. The reduction in workload and time-losses that arise from ICS adoption increase with distance to firewood collection sites. Therefore, we expect the ICS adoption rate to be higher among households largest that live further away from fire wood collection sites, compared to households that live relatively closer to such wood collection points.

We expect to find a negative association between the following variable and the ICS adoption rate:

- Distance to closest stove distribution point. Since transport remains an issue, households that would have to travel long distances to bring their stoves back home are expected to be less likely to buy a stove. Therefore, we expect the ICS adoption rate to be higher among households that live close to a stove distribution point, compared to households that live further away from such a distribution point.

The expected association between the following variable and the ICS adoption rate is less straightforward:

- Household headship. We expect the ICS adoption rate to be higher among single-headed households, compared to couple-headed households. Single-headed households are often women-headed households. Household headship is thus strongly linked to the gender of the household head. Women generally tend to have stronger preferences for ICSs than men. They are expected to adopt a stove when they have the authority to make such a purchase decision as a household head. As a result, ICS adoption is expected to be higher among women-headed households compared to men-headed households.

5.8 Data collection

This sub-section discusses different aspects of the data collection process of the study. First, the sub-section describes the data sources that will be used to assess the effects of the RCT treatment conditions. Subsequently, the execution of the endline survey is discussed in more detail. Amongst others, the selection and training process of the enumerators as well as other data collection procedures are described.

5.8.1 Data sources

Quantitative and qualitative data will be collected in order to gain a thorough understanding on the (short term) effects of the treatment conditions. The household endline survey is the most important data source for this study. The results of the endline survey will be supported by information gathered through focus group discussions and air quality measurements.

Household endline survey

The endline survey will be conducted after approximately one year of implementation of the ICS intervention. This survey will be a collaborative effort of the WUR, NLRC and ERCS and will also function as endline survey for the 3FM pneumonia project as whole. Therefore, the survey will cover many topics, just like the baseline survey. However, unlike the baseline survey, the endline survey will cover the impact indicators of this stove intervention study in great detail. Amongst others, questions on ICS adoption and utilisation, firewood consumption, time allocation patterns and child health and mortality, will be included in the survey. As mentioned before, the endline survey will cover 264 target villages. In each village, approximately three households will be randomly selected from the target population for survey participation. We thus aim to conduct a total of 792 household endline surveys.

Air quality measurements

Indoor air quality measurements will be carried out simultaneously with the endline survey. We will measure the static CO and $PM_{2.5}$ concentrations in kitchens by the use of PM sensors and CO data loggers. In addition, we will measure the exposure of under-fives to (indoor) air pollution with passive diffusion tubes. These measurements will be carried out in a subset of the households that will participate in the endline survey. We will randomly select one out of the three participating households per village.

Focus Group Discussions

Focus Group Discussions (FGDs) will be conducted at different stages of the study. These FGDs will be conducted at the start, halfway and at the end of the study period. For these FGDs, a few target sub-districts will be randomly selected. Subsequently, a specific number of villages will be selected randomly within these selected sub-districts. In these selected villages, FGDs will be carried out with target households, RC volunteers, health professionals and other influential persons. These discussions will cover the ICS adoption process, cooking practices and child health. We will use the qualitative data to interpret and elaborate on the quantitative results. The qualitative component of this study enables us to explore how the ICS intervention affects the lives of target households and to thereby identify underlying processes of change.

5.8.2 Data collection procedures

Household endline survey

The endline survey as well as the FGDs will be conducted by local RC volunteers. In seventeen from the twenty target sub-districts, five RC volunteers are active. In the remaining sub-districts, approximately ten RC volunteers are active. These volunteers will receive training on survey technique and the selection procedure of participants. Furthermore, they will be trained on protocols to maintain the confidentiality of participants' answers to the extent possible in complex field situations.

Following the survey protocol, the RC volunteers will clearly identify themselves as independent enumerators (both verbally as well as visually). They will emphasize that, although this survey covers a RC project, survey responses will be received and analysed by an independent research organisation. To further reduce the risk of enumerators being associated with the RC, volunteers will conduct surveys in villages where they are not active as a RC volunteer. These precautionary measures will minimise response bias as well as potential tension between enumerators, communities and the RC.

The survey data will be collected electronically by the use of KoBoToolbox software. Immediately after collection, the data will be uploaded to a web-based database (Kobo Server). The core research team, existing out of WUR researchers and local research assistants, will monitor the execution of the endline survey.

Air quality measurements

The IAQ measurements will be conducted by students from a, at the moment not yet specified, local university. They will receive extensive training on the usage of air measurement equipment and data collection procedures. We are planning to use PM sensors, CO data loggers and passive diffusion tubes to measure IAP. We will select these instruments during the study period. The passive diffusion tubes will be read by the field worker at the time of collection. These tubes will be read a second time, blind to the first reading, by a member of the core research team. The PM sensors and CO data loggers contain dataloggers. This enables them to store minute-by-minute data over the entire measurement period. After monitoring, the resulting dataset will be downloaded into a computer and uploaded in the shared Dropbox folder.

Focus Group Discussions

For the FGDs, an interview guide will be prepared. This guide will contain all the questions that we want to discuss during the FGD as well as clarifying and probing questions that can be used by the interviewer to guide the discussion. A rapporteur will take notes of the discussion along with his or her observations. These notes will be summarized after the discussion and all forms concerning the FGDs will be uploaded in a shared Dropbox folder.

5.9 Ethical considerations

Random assignment of treatment conditions

Excluding target households in the intervention area from the full treatment may be perceived as unethical. However, we need to vary the (levels of) treatment of the target households in order to be able to research the effects of the treatment factors. As a result, a share of the target households will not receive the full stove subsidy, another share of households will not receive complete information on the benefits of the *Mirt* stove and another share of households will receive neither. Consequently, these target households are expected to be less motivated to buy a *Mirt* stove.

However, we do not deprive these households from the possibility to purchase a *Mirt* stove. Through the 3FM pneumonia project, these households will receive (limited) information on the *Mirt* stove and they will get access to this stove for cost-covering prices. Furthermore, it's worth noting that varying levels of treatment is an inherent feature of RCTs. Short-term restrictions in RCT treatments serve the greater good. After all, we conduct this RCT to inform effective strategies to maximise the adoption of improved cookstoves. The results of this study can contribute to effective strategies for improving living conditions and respiratory health of under-fives in Ethiopia.

Ethical clearance and informed consent

We will apply for ethical clearance for this study from the Ethics Committee of the WUR as well as from the ethics Committee of a, at the moment not yet specified, local university. We already obtained ethical clearance from the (regional) administrative offices for the study as a whole as well as for the stove introduction meetings and household survey in particular. Additionally, ethical clearance will be obtained from local health bureaus for the endline household survey.

Returning to the subject of RCT participation, we will not obtain informed consent of households that will participate in this trial. Since obtaining informed consent for the RCT can have substantial effects upon the outcomes of the trial, we argue that there should not be a need to obtain informed consent from participating households (Verheggen & Wijmen, 1996). Nevertheless, households will be informed about the fact that they are part of a study. In addition, we will obtain informed consent from all survey participants. Prior to each data collection session, the enumerators will introduce themselves, explain the purpose of their visit and obtain informed consent of the respondent. Informed consent will be obtained for IAQ measurements as well.

5.10 Validity of study

5.10.1 Internal validity

Internal validity refers to the confidence that we can place in the evidence that this study provides on the causal relationship between the treatment factors and the selected outcome variables.

Confounding factors

Factors that influence both the dependent and independent variables are a threat to the validity of causal inferences made by a study. The implementation of other activities that are part of the 3FM pneumonia project might be such a confounding factor. These other project activities might affect the outcomes of the ICS intervention, since they also focus on creating awareness on (causes of) pneumonia. We expect the confounding effects of this factor to be relatively small since these other standardized project activities are implemented in all study sub-districts.

History

Certain large-scale events that happen outside of this study, such as the implementation of other ICS interventions or projects, can affect the study results. If this is the case, it becomes impossible to determine whether change in the outcome variables is due to the ICS intervention or the other large-scale event. After some research, we concluded that no other large-scale ICS interventions will be implemented during the coming year in our study areas.

Study design

As discussed before, no adequate baseline data was collected. Taking into account the time and budget constraints, we decided to not conduct a new baseline survey. This impact study will thus make use of endline data solely, which can be risky.

Our primary outcome variable is the adoption the *Mirt* stove and our secondary outcome variables cover *Mirt* stove utilisation rate, firewood consumption, time allocation patterns of women and children, indoor air quality and reported child health. While the primary outcome variables are directly influenced by the treatment factors, the secondary outcome variables are expected to be only indirectly affected by the treatment factors through ICS adoption and usage. At the same time, a broad variety of other (external) factors are also expected to affect these secondary outcome variables. As a result, it will be quite challenging to detect short-term effects of the treatment factors on the secondary outcome variables.

The power calculations indicated that it will be particularly challenging to detect the treatment effects on child health (See sub-section 5.5). It certainly does not help that no significant health effects of *Mirt* stove usage were found until now. It is not yet clear if the health benefits of the *Mirt* stove are relatively small or if its health improving potential is attenuated through, for example, continued use of traditional stoves alongside the *Mirt* stove. Either way, it will be relatively hard to detect the impact of the treatment factors on child health.

Furthermore, it is important to bear in mind that some of the secondary outcome variables are very challenging to measure. While fuel consumption and time allocation patterns can be assessed relatively easy, it is very costly and time consuming to measure IAQ. The use of objective and robust measurements of child health is even more challenging. This should be taken into account when the data is analysed and conclusions are drawn.

Sources of bias

Furthermore, the internal validity of this study is threatened by several potential sources of bias. Amongst others, response bias can occur. It is quite likely that respondents will consciously or subconsciously give responses of which they think that they are (socially) desirable or expected. Although precautionary measures are put in place, there is still a chance that respondents link the endline survey to the RC. If they do so, their responses can be influenced by their dependencies, associations and former experiences with (the work of) the RC. In that case their answers do not accurately reflect reality.

Another way in which response bias can occur, is when subjects learn that they have been excluded from full treatment. As a result, they may experience feelings of resentment and demoralisation, which could motivate them to give inaccurate answers on the questions that they are being asked. This, again, underlines the importance of a carefully and discreetly executed trial.

Furthermore, spill-over effects can also introduce bias. We aim to prevent this by selecting treatment factors that do not easily spill over to non-targeted households. The stove subsidy coupons as well as the invitations for the introduction are personalised and cannot be passed onto others. After the stove introduction meetings took place, it is plausible that stove adopters share information and personal experiences with the *Mirt* stove with others. However, this is not perceived to be a spill-over of treatment conditions but a result from social learning processes. Since we are interested in measuring social learning effects, we will include questions in the end-line survey to capture these effects.

As the discussed potential threats of internal validity are considered to not substantially affect the results of this study, the internal validity of the study is assessed to be relatively high.

5.10.2 External validity

The external validity is the extent to which the findings of this study apply to other contexts.

Selection process

The external validity of this study is negatively affected by the selection process of the study villages. Our study sub-districts are the sub-districts that were selected for the 3FM pneumonia project. The RC has selected these sub-districts based on their vulnerable status. In most of these vulnerable areas, the RC has implemented various other projects in the past. In a certain way, the 3FM pneumonia project thus builds further on the impact realised by these previous projects. Amongst others, these previous projects of the RC are expected to have improved living conditions and child health. Furthermore, through participating in these projects, people have learnt about the work of the RC and how it can impact their life. This is expected to positively affect their trust in the RC, which increases the uptake and impact of new projects such as this 3FM pneumonia project. As a result, we cannot assume that the situation in our study villages represent the situation in an average Ethiopian village or that this stove intervention would have similar impact in other Ethiopian villages that were not covered by similar development projects in the past. Therefore, we need to be careful with generalising the findings of this study.

Research population

The target population of the 3FM pneumonia project consists of households with under-fives. Therefore, the ICS intervention will also focus on this specific group of households. However, in the design phase of this trial, it was decided by the RC that the most vulnerable households with under-fives should be excluded from trial participation. These households should receive a free stove, regardless of the treatment condition that will be assigned to their village, due to their vulnerability status. Consequently, this subset of households is not part of the research population and the findings on the treatment factors cannot be generalized to these households without further research.

3FM pneumonia project

A broad set of project activities will be implemented next to the ICS intervention. Synergy between the ICS intervention and these other project activities might bias the results of this study. It is possible that the treatment factors are less effective in increasing ICS adoption when they are implemented separately from the other project activities that raise awareness on (causes of) pneumonia. Therefore, we need to be careful with generalising the results of this study to contexts in which these enabling factors are not present.

Hawthorne effects

Furthermore, the external validity of this study can be threatened by the Hawthorne effect. This refers to an effect that simply originates from the fact that subjects know that they are being studied. We strive to inform households from the study villages in an appropriate way about their participation in this study and at the same time also prevent such reactive effects. Therefore, we strongly emphasise the link between the ICS intervention and other project activities in order to minimise the extent to which the provided information influences people's state of mind.

All in all, the external validity of this trial is rather limited. This RCT is conducted in a highly specific context. Consequently, we should be cautious in generalizing the results of this study (to other parts of Ethiopia). Additional intervention-based studies in other populations need to be implemented to verify the results of this study.

5.11 Timeline

The timeline of study procedures is as follows:

• **July 2017.** Conducting the baseline study for the 3FM pneumonia project.

• **January 2018 - June 2018.** Writing the study proposal; meetings with NLRC and ERCS staff. Applying for approval from ERCS and respective (regional) administrative offices for the execution of an explorative research and the entire study.

• **August 2018.** Planning and preparing an exploratory mission to Ethiopia. Developing and testing the expert survey and the FGDs.

• **September 2018 - October 2018.** Mission to Ethiopia. Conducting FGDs with representatives from ministries, Non-Governmental Organisations (NGOs), health professionals, local influential persons, target households and RC staff. Gathering data from local health facility records. Exploring the possibilities for the design of the study.

• **November 2018 - December 2018.** Consensus building around the project design. Writing implementation guidelines and randomization of treatment conditions.

• **March 2019 - April 2019.** Implementation of the ICS intervention in line with the RCT design.

• **June 2019.** Qualitative data collection. Assessment of the total number of *Mirt* stoves produced and sold. Conducting FGDs to evaluate the current situation in the study villages and to explore the channels through which the ICS intervention affects the practices and living conditions of target households.

• **November 2019.** Again, monitoring the general situation in the study villages.

• **February 2020 - March 2020.** Preparations for the endline survey. Pilot study. Testing of the quantitative survey instrument; in-depth interviews on participants' understanding of the survey instrument.

• **April 2020.** Endline study. Conducting the household endline survey as well as the FGDs to illustrate and clarify the survey results. Measuring IAQ and exposure of under-fives to IAP. Conducting data analyses, assessing the effects of the ICS intervention and writing a final report on (the results of) this study.

• **April 2020 - September 2020.** The ICS intervention will be implemented in the other sub-districts to which the 3FM pneumonia project expanded throughout the last year.

6. Analysis of experts' beliefs

This section discusses the expert survey that was conducted to elicit experts' beliefs about the efficacy of the ICS intervention. To be more specific, experts were asked to forecast future research findings on the effectiveness of the free distribution of ICSs. The free distribution of ICSs is one of the treatment factors of the 2x2 factorial design of the ICS intervention that was discussed in the previous section (See table 5). Basically, respondents were asked to estimate the difference between the impact of Treatment 1 and 3 (stove introduction meeting without access to a free stove) and Treatment 2 and 4 (stove introduction meeting with access to a free stove). Subsequently, this section will specify the methodology of the analysis of the elicited experts' beliefs. The objective of this analysis is to evaluate the reported beliefs and comparing these beliefs to the literature-based predictions. The section finishes with the empirical results and the limitations of this analysis.

This section thus presents the results for RQ4: What are experts' beliefs on the impact of this ICS intervention and how do these predictions compare to results from previous studies on comparable interventions?. Although it is important to bear in mind that the respondents are asked to forecast the impact of one of the treatment factors of the RCT on the ICS intervention in specific. The other treatment factors of the RCT are excluded from this analysis. Therefore, the results of this analysis will give a rough indication of experts' beliefs on the impact of the treatment existing out of the distribution of a free *Mirt* stove. If predictions differ greatly among the surveyed experts, it would be interesting for the RC to explore the source of variation in predictions.

After the finalization of the ICS intervention, elicited experts' beliefs can be compared to the outcomes of this intervention to assess the actual accuracy of their beliefs.

6.1 Expert Survey

Before the survey starts, respondents are introduced to the ICS intervention and the specific treatment of which they will be asked to forecast the effectiveness. Next, the context of this ICS intervention is described. Accordingly, the 3FM pneumonia project in Ethiopia and the study design are introduced.

The survey starts with some introductory questions and continues with an extensive set of prediction questions. The last set of questions focuses on the predictions of the respondents regarding the research findings on the effectiveness of the specified treatment factor. The selected outcome variables are individually introduced (See table 10). This set of outcome variables is not identical to the set of outcome variables that was selected for the impact study. For each outcome variable, the related hypothesis and the indicator are shared. Respondents are asked to predict the treatment effect on the specific outcome variable that would be found one year after implementation of the intervention. They can choose from the following options to express their prediction: Large negative effect (--), small negative effect (-), no significant effect (0), small positive effect (+) and large positive effect (++). Respondents are also asked to rate their confidence for each provided prediction on a 3-point scale.

6.2 Methodology

This sub-section discusses the methodology used to analyse experts' beliefs. The sub-section first discusses the study sample, continues with the hypotheses, then clarifies the strategy for assessing the forecast accuracy and finishes with a description of the empirical strategy used for data analysis.

Sample

Staff members of the ERCS, NLRC and Zambia Red Cross Society (ZRCS) were invited to participate in the expert survey during the period from September to November 2018. The participating RC staff members have varying positions within these RC societies. Field officers from the ERCS to Planning, Monitor and Evaluation officers from the NLRC are represented in the dataset. Furthermore, two RC staff members from the ZRCS have participated in the survey. These staff members are working on the implementation of the 3FM pneumonia project in Zambia.

In addition, a group of master students from WUR were invited to take part in the expert survey in January 2019. At that moment, these master students participated in a course called "Impact assessment of policies and programmes". These students were comprehensively informed on the 3FM Pneumonia project and the Ethiopian context before they were asked to participate in the expert survey.

The following table presents the number of respondents per group of experts:

Table 8. *Expert survey – Number of respondents.*

Group of experts	Number of respondents
NLRC staff members	**15**
Staff members from other RC national societies	**14**
ERCS	*12*
ZRCS	*2*
WUR students	**33**

Note: This table lists the number of respondents of the expert survey for the three groups of experts. The number of respondents from the expert group "Staff members of other RC national societies" is broken down by national society.

Hypotheses

I will test the five hypotheses stated below. These hypotheses were formulated based on general expectations of involved researchers.

H1: Average prior beliefs and forecast accuracy differ across groups of experts.

H2: Average prior beliefs are positively correlated with the level of personal support for the ICS intervention.

H3: The accuracy of a prediction is positively correlated with the reported level of confidence.

H4: Forecast accuracy is positively correlated with the level of familiarity with the context of the ICS intervention.

H5: Forecast accuracy is positively correlated with the level of familiarity with impact studies.

Forecast accuracy

I will assess the forecast accuracy by comparing individual forecasts with the literature-based predictions. In order to do so, I have first translated the effects that were found by previous studies on comparable ICS interventions to effect sizes. The resulting effect sizes are used to categorise the effects found in the literature in identical categories as from which experts could choose to express their predictions. Unfortunately, not all outcome variables from the expert survey are covered by previous studies. As a result, the accuracy of predictions regarding these outcome variables cannot be assessed.

I have used different techniques to classify the effects found by previous studies of ICS interventions on continuous and binary outcome variables. For continuous variables, I have expressed the size of treatment effects in standard deviation units (sdu's). The classification in the table below was used to classify the resulting effect sizes.

Table 9. *Expert survey - Classification of effect size in sdu's.*

Treatment effect size in sdu's	Classification
≤ 0.05	No significant effect
0.05 < x ≤ 0.30	Small effect
> 0.30	Large effect

Note: This table specifies how I have classified the effects of ICS interventions found in the literature on continuous variables to enable the comparison between the expert beliefs and the literature-based predictions. These treatment effects are classified into three categories based on their size expressed in standard deviations units.

Effects found by previous studies on binary outcome variables cannot be expressed in sdu's. These effects are classified based on the conclusions from the studies in question. For these variables, qualitative criteria are thus used to classify the effect sizes into the same categories as mentioned in table 9.

Table 10. *Expert survey – Treatment effects found in the literature.*

Question *Survey*	Outcome variable *Survey*	Effect size *Literature*	Reference *Literature*
1	Knowledge on cooking practices	-	-
2	Knowledge on ventilation practices	-	-
3	Utilisation rate of ICSs	Large	Bensch & Peters (2015)
4	Kitchen ventilation	-	-
5	Cooking time	Large	Bensch & Peters (2015)
6	Outside cooking	Large	Bensch & Peters (2015)
7	Firewood consumption	Large	Bluffstone et al. (2017)
8	Choice of cooking fuel	-	-
9	Protection of children against smoke	-	-
10	Indoor air quality	Large	Bruce et al. (2004); Dutta et al. (2007); Masera et al. (2007); Chengappa et al. (2007)
11	Exposure of under-fives to IAP	Large	Bruce et al. (2004)
12	Child respiratory health[1]	Not significant	Mortimer et al. (2017); Smith et al. (2011)
13	Child respiratory health[2]	-	-
14	Overall child health[3]	Not significant	Mortimer et al. (2017)

Note: This table gives an overview of the outcome variables regarding which the experts were asked to predict the treatment effect of free distribution of *Mirt* stoves (column 2). This table also classifies the effect sizes on these outcome variables found by previous studies (Column 3) as well as the authors of the studies in question (Column 4).

[1] Indicator: self-reported prevalence rate of ARI among under-fives

[2] Indicator: prevalence rate of ARI among under-fives reported by local health facilities

[3] Indicator: ARI-related mortality rate among under-fives

Table 10 presents the results of this process. A prediction of a respondent on the treatment effect on one of the outcome variables is perceived to be accurate when it matches the literature-based prediction. Respondents will receive one point for each accurate prediction, with a maximum score of 8. Their total score will indicate the overall accuracy of their predictions, their so-called forecast accuracy.

Empirical strategy

For Hypothesis 1, I will first compare the mean predicted effect size across the different groups of experts. For this purpose, I will create a variable on average predicted effect size and compare the means of the three groups of experts. I will also assess the accuracy of predictions of the three groups of experts. In order to do so, I will first construct a graph with point estimates and 95% confidence intervals for the predictions regarding each of the 8 outcome variables, grouped by expert group. Next, I will create a variable on the overall accuracy of predictions and compare the mean forecast accuracy of the three groups of experts.

Furthermore, I will run several additional descriptive analyses. For hypothesis 2, I will investigate the correlation between the average prior beliefs and the level of personal support for the ICS intervention. Subsequently, I will estimate the correlations between the accuracy of each separate prediction and the reported level of confidence to test hypothesis 3. To conclude, I will assess the associations between forecast accuracy and multiple personal characteristics of respondents to test hypotheses 4 and 5.

6.3 Empirical results

This sub-section first presents the descriptive statistics. Whereupon, the empirical results of this analysis are presented and discussed by hypothesis.

Descriptive statistics

The table below presents the descriptive statistics for the set of variables that will be included in this data analysis.

Table 11. *Expert survey – Descriptive statistics.*

Variable	Description	Obs.	Mean	Std. Dev.	Min	Max
Expert group	Groups of experts **[**0=NLRC staff, 1=Staff other RC national societies, 2=WUR students**]**	62	0.53	0.50	0	1
Perceived importance	Personal opinion on the importance of the ICS intervention **[**ten-point scale**]**	60	1.48	0.91	-1	2
Fam. project context	Familiarity with the context of the project **[**two-point scale**]**	62	0.53	0.33	0	1
Fam. impact studies	Familiarity with the context of the project **[**ten-point scale**]**	62	4.58	1.99	1	10
Effectsize_*Nr. of outcome variable*	Predicted treatment effect on outcome variable **[**-2=Large negative, -1=Small negative, 0=Not significant, 1=Small positive, 2=Large positive**]**	62	-	-	-2	2
Confidence_*Nr. of outcome variable*	Confidence in specific prediction **[**0=Not at all confident, 1=Somewhat confident, 2=Very confident**]**	62	-	-	0	2
Mean effect	Mean of all predicted treatment effects	62	1.00	0.50	0.07	2
Accuracy prediction_ *Nr. of outcome variable*	Indication of accuracy of specific prediction **[**0=Prediction is accurate, 1= Prediction is inaccurate**]**	62	-	-	0	1
Sum accuracy	Number of accurate estimations	62	2.44	1.90	0	6

Note: This table describes the summary descriptives for all the variables that are included in the analysis of the experts' beliefs. For each outcome variable, two variables are included in this analysis that indicate predicted effect size and reported level of confidence (*Effect size* and *Confidence*). For eight outcome variables, an indicator of the accuracy of experts' predictions is included in the analysis (*Accuracy prediction*).

Hypothesis 1

The level and accuracy of predictions are hypothesised to differ across groups of experts ($P<0.05$). In order to test this hypothesis, first the predicted effect sizes are compared across the three groups of experts.

The results of a one-way analysis of variance (ANOVA) indicate that the mean predicted effect size do significantly differ across the groups of experts ($F(2,59)=16.44$, $p<0.001$). A Tukey post-hoc test revealed that the mean predicted effect sizes among WUR students (0.77 ± 0.35, $p<0.001$) and staff members of the NLRC (1.05 ± 0.51, $p=0.007$) are significantly lower than the mean predicted effect size among staff members of the other RC national societies (1.53 ± 0.39). Prior beliefs about the impact of the specified treatment factor are thus most optimistic amongst staff members of the ERCS and ZRCS.

Additionally, the forecast accuracy is hypothesised to differ across expert groups ($P<0.05$). To test this, first the accuracy of the predictions regarding each of the eight outcome variables are compared.

The figures below show the mean predictions and the 95% confidence intervals for the eight outcome variables, grouped by expert group. In these figures, the horizontal red lines represent the literature-based predictions. These graphs indicate that, regarding most of the outcome variables, the predictions of staff members of the ERCS and ZRCS are on average most accurate. While WUR students are found to be least accurate in their predictions regarding the treatment effects, with the exception of their predictions regarding the treatment effects on child respiratory health and overall child health. Treatment effects on child health are expected to be rather small, since the intervention only indirectly affect the respiratory health in children. Also, not all factors that influence the respiratory health of children are targeted by the intervention. Remarkably, a large share of the respondents have overestimated the impact of the intervention on child health. The share of respondents that have overestimated the impact of the intervention on child health appears to be higher among the RC staff members than among the WUR students.

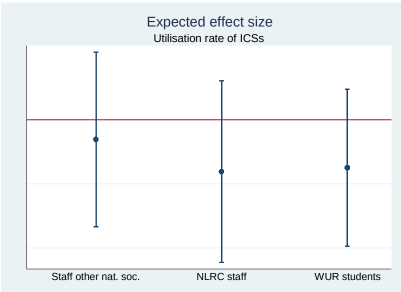

Figure 3: Experts' predictions - Utilisation ICSs

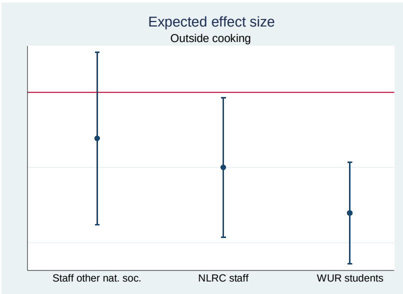

Figure 4: Experts' predictions – Cooking time

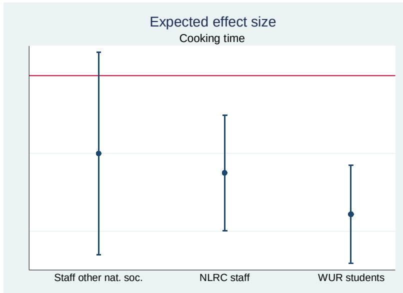

Figure 5: Experts' predictions – Outside cooking

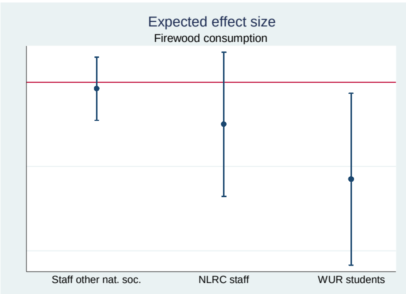

Figure 6: Experts' predictions – Fuel consumption

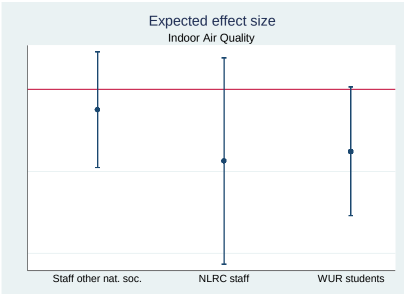

Figure 7: Experts' predictions – IAQ

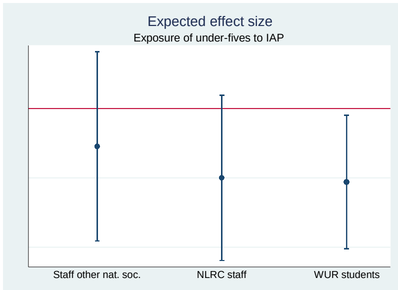

Figure 8: Experts' predictions – Exposure IAP

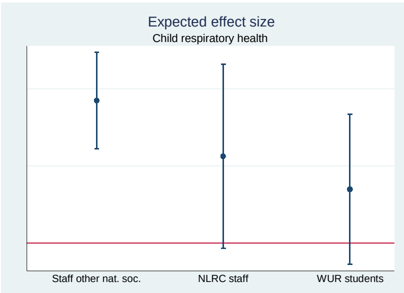

Figure 9: Experts' predictions – Child respiratory health

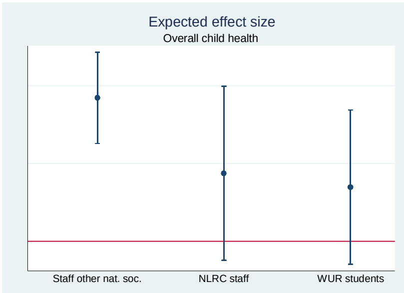

Figure 10: Experts' predictions – Overall child health

Note: These figures show plots of the predicted treatment effects regarding eight outcome variables, grouped by expert group (Staff other RC national societies, NLRC staff and WUR students). These figures include mean predicted effect sizes (dots), 95% confidence intervals (vertical blue lines) and literature-based predictions (horizontal red lines).

Moreover, the overall accuracy of the predictions of these three groups of experts is compared. The overall accuracy of respondents' predictions, their forecast accuracy, is indicated by the number of accurate predictions provided. The results of a Tukey post-hoc test indicate that staff members of the ERCS and ZRCS (4.15 ± 1.99) are, on average, more accurate in their estimations regarding treatment effects than WUR students (1.64 ± 1.27, $p<0.001$). Their forecast accuracy does not differ significantly from the forecast accuracy of the NLRC staff members (2.69 ± 2.02, $p=0.053$). These results confirm that forecast accuracy differs significantly across expert groups.

Hypothesis 2

The level of personal support for the ICS intervention is hypothesised to be positively associated with the mean predicted effect size (P<0.05). Table 12 shows the correlation coefficients that express the association between these two variables for each of the three expert groups. The correlation coefficients strongly differ across the expert groups and are all highly insignificant. These results do not support Hypothesis 2. No proof for a significant positive association between the level of personal support for the ICS intervention and mean predicted effect size has been found.

Table 12. *Hypothesis 2 – Spearman's Correlation.*

	Personal support (*Staff other RC national societies*)	Personal support (*NLRC staff*)	Personal support (*WUR students*)
Predicted effect size (*Staff other RC national societies*)	-0.483		
Predicted effect size (*NLRC staff*)		0.384	
Predicted effect size (*WUR students*)			0.086

Note: This table shows the correlation coefficients between the level of personal support for the ICS intervention of the RC and the mean predicted treatment effect, grouped by expert group.
Correlation coefficients are standardized. * p<0.10, **p<0.05, *** p<0.01

Hypothesis 3

The accuracy of respondents' predictions is hypothesised to be positively correlated with their reported confidence levels (P<0.05). In other words, the measure of confidence is expected to be predictive of forecast accuracy. Table 13 shows the correlations between the accuracy of the individual predictions and the reported confidence in these specific predictions. The accuracy of respondents' predictions regarding the impact of the specified treatment factor on the utilisation rate of ICSs (3), cooking time (5), cooking location (6), firewood consumption (7), IAQ (10) and exposure of under-fives to IAP (11) are positively associated with their confidence in these predictions (P<0.01). At the same time, a positive correlation is found between reported confidence levels and the predicted effect size. It appears that respondents who are less confident about their predictions, have the tendency to underestimate the effects of the intervention.

Although no causal claims can be made, the most obvious interpretation of these positive correlations would be that a prediction is more likely to be accurate when a respondent indicates to be confident about this specific prediction.

The associations between the level of confidence and accuracy of predictions on the impact of the specified treatment factor on child respiratory health (12) and overall child health (14) are found to be an exception to this phenomenon. These variables are found to be negatively correlated. However, these coefficients are relatively small and not statistically significant. If we take a closer look to the predictions on the impact of the specified treatment factor on these two outcome variables, it becomes clear that a large proportion of respondents have overestimated the impact on child health. At the same time, these respondents have indicated that they are (relatively) confident about these predictions. This error of judgement could explain the sudden change in sign of correlation between reported confidence levels and accuracy of predictions.

Table 13. *Hypothesis 3 – Spearman's Correlation.*

	Accuracy 3	Accuracy 5	Accuracy 6	Accuracy 7	Accuracy10	Accuracy11	Accuracy12	Accuracy14
Confidence3	0.623***							
Confidence5		0.488***						
Confidence6			0.484***					
Confidence7				0.605***				
Confidence10					0.626***			
Confidence11						0.536***		
Confidence12							-0.141	
Confidence14								-0.166

Note: This table shows the correlation coefficients between the accuracy of predictions on the impact of the specified treatment factor on eight outcome variables and the reported levels of confidence in these individual predictions.
Correlation coefficients are standardized. * p<0.10, **p<0.05, *** p<0.01

Hypothesis 4

Forecast accuracy is hypothesised to be positively correlated with the level of familiarity with the context of the ICS intervention (P<0.05). The correlation coefficients between these two variables strongly differ across the three groups of experts. Forecast accuracy is found to be significantly and positively correlated with level of familiarity with the context of the ICS intervention among NLRC staff members and staff members of other RC national societies. For these groups of experts, their level of familiarity with the context of the ICS intervention appears to be predictive of their forecast accuracy.

Table 14. *Hypothesis 4 – Spearman's Correlation.*

	Familiarity context (*Staff other RC national societies*)	Familiarity context (*NLRC staff*)	Familiarity context (*WUR students*)
Forecast accuracy (*Staff other RC national societies*)	0.652**		
Forecast accuracy (*NLRC staff*)		0.552**	
Forecast accuracy (*WUR students*)			-0.115

Note: This table shows the correlation coefficients between the level of familiarity with the context of the ICS intervention of the RC and the overall accuracy of predictions regarding the treatment effects, grouped by expert group.
Correlation coefficients are standardized. * p<0.10, **p<0.05, *** p<0.01

Hypothesis 5

Forecast accuracy is hypothesised to be positively correlated with the level of familiarity with impact studies (P<0.05). The correlation coefficients substantially differ across the expert groups and are all insignificant. No proof for a significant positive association between respondents' level of familiarity with impact studies and forecast accuracy has been found.

Table 15. *Hypothesis 5 – Spearman's Correlation.*

	Familiarity studies (*Staff other RC national societies*)	Familiarity studies (*NLRC staff*)	Familiarity studies (*WUR students*)
Forecast accuracy (*Staff other RC national societies*)	-0.117		
Forecast accuracy (*NLRC staff*)		0.362	
Forecast accuracy (*WUR students*)			0.094

Note: This table shows the correlation coefficients between the level of familiarity with impact studies and the overall accuracy of predictions regarding the treatment effects, grouped by expert group.
Correlation coefficients are standardized. * p<0.10, **p<0.05, *** p<0.01

6.4 Limitations

This sub-section discusses the limitations of the expert survey and the analysis of the collected data. As mentioned before, the aim of this expert survey is to assess experts' beliefs regarding the study findings on the impact of the distribution of free ICSs. As a result, this analysis is limited to assessing experts' beliefs concerning this treatment factor and comparing a subset of the treatment conditions of the RCT. The other treatment factors or treatment conditions that are part of the RCT (See sub-section 5.1) are excluded from this analysis.

Furthermore, the extent to which we can assess the accuracy of reported predictions is limited due to the fact that the impact of ICS interventions is still unexplored regarding a large share of the outcome variables that are included in the expert survey. As a result, only the accuracy of reported predictions regarding 8 out of the 14 outcome variables can be assessed based on results from previous studies.

Another important limitation of this analysis concerns the small sample size. The sample size was even further reduced by missing data. The missing data problem was mainly caused by survey questions that were left unanswered. A relatively small sample size is generally associated with low statistical power. It is possible that, based on our sample, there is not sufficient evidence to detect a significant correlation. Consequently, the small sample size is an important limitation of this analysis due to its negative effect on power.

To conclude, a few remarks have to be made about the quality of the data collected. Firstly, language can pose a barrier in the collection of high-quality data. Since the survey is developed in English, language barriers could have occurred for respondents that have poor English skills. This issue could explain the relatively large variance in predictions of staff members from the ERCS and ZRCS. Some of these staff members have provided seemingly unrealistic predictions.

Although, the latter phenomenon could also be caused by the social desirability bias. This bias refers to the tendency of survey respondents to give (socially) desirable or expected responses instead of choosing responses that reflect their true feelings. For example, RC staff members could feel obliged to take a positive standpoint with regard to the work of the RC. In this case, they would report very high predictions regarding the impact of the ICS intervention, regardless of the outcome variable. Although I tried to limit social desirability bias by assuring privacy and confidentiality, it is important to bear in mind that the answers of respondents might be biased by social desirability.

Lastly, the way in which I have set up the questions and answer options can also create bias and affect data quality. Although the survey was set up as neutral as possible, it does include our hypotheses on the relationships between the specified treatment and the selected outcome variables. This is done for clarification purposes. However, mentioning these hypotheses could also influence the answers of the respondents. Only a few respondents have reported expectations that contradict with the sign of hypothesised relationships.

7. Concluding remarks

In Ethiopia, pneumonia is still the leading single disease killing under-fives. The prevalence rate of childhood pneumonia within the study sample is 19.7 percent. In general, the knowledge of health professionals and households in rural parts of Ethiopia on pneumonia is very limited. During the previous years, training of health professionals and an intensive BCC plan have been implemented as part of the 3FM pneumonia project to improve awareness and knowledge on pneumonia. Amongst others, risk factors, prevention measures and advised treatment of childhood pneumonia have been extensively discussed.

Indoor air pollution is one of the established risk factors of childhood pneumonia that is perceived to contribute substantially to the burden of childhood pneumonia in Ethiopia. However, a large majority of the respondents of the baseline survey appeared to be not aware of the health effects of IAP. Since traditional cooking practices are a major cause of dangerous IAP, the RC also aims at promoting improved cooking practices through this pneumonia reduction project. Amongst others, an ICS intervention will be implemented to promote the locally produced *Mirt* stove. Despite the fact that *Mirt* stove usage has been heavily promoted during the last decades by the Ethiopian government as well as various NGOs and organisations, only four percent of rural Ethiopian households currently own this fuel-efficient cookstove.

This document includes the experimental design of an impact study that will evaluate this ICS intervention. The primary objective of this RCT is to assess how effective stove subsidies and information provision on stove benefits are in increasing ICS adoption. By means of this trial, we will study the effects of subsidies and information provision on *Mirt* stove adoption and usage as well as economic and health outcomes.

Since women generally tend to have stronger preference for ICSs than men, we aim at motivating men to support their wives in this adoption decision through information provision on *Mirt* stove benefits. We will vary the content of the stove introduction meetings to assess the impact of information provision on stove adoption through the intra-household bargaining processes. We hypothesize that men value economic benefits higher than health benefits since they would directly experience these benefits. Therefore, half of the meetings will focuses primarily on the expected health benefits of the *Mirt* stove. At the other meetings, both the expected health and economic benefits of the *Mirt* stove will be discussed.

Furthermore, we aim at reducing the liquidity- and credit constraints of rural Ethiopian households through stove subsidies. Four different vouchers will be issued at the stove introduction meetings, representing 0%, 40%, 70% and 100% subsidies. The results on this treatment factor will inform us on the optimal level of subsidy, taking into account the adoption rates and sustainability of stove usage at different subsidy levels.

With these two treatment factors, eight different treatment conditions can be created. The resulting treatment conditions will be randomized on village level. Approximately 483 villages will be included in this randomization process. Since the 3FM pneumonia project focuses on improving the health of under-fives, the study population of this trial exists out of households with under-fives. A total of 10.736 households with at least one child under-five years of age live in the study villages. However, on request of the RC the 10 percent most vulnerable households with under-fives will be excluded from trial participation. These households will receive a free *Mirt* stove, regardless of the treatment condition that will be assigned to their village.

Based on the power calculations, we expect to be able to detect the main effects of the free distribution of ICSs and information provision on ICS adoption, utilisation rate of ICSs, household cooking efficiency and (exposure to) IAP. In line with our expectations, it will be relatively hard to detect the impact of the specified treatment factors, through ICS adoption, on child health. While we might be able to detect the treatment effects on childhood pneumonia, this will most certainly not be the case for pneumonia-specific mortality among under-fives.

The expert survey has informed us on the beliefs of the NLRC staff members, staff members of other RC national societies and WUR students on the impact of the free distribution of ICSs. Amongst others, staff members of the ERCS and ZRCS are found to be most optimistic about the impact of this treatment factor. While the WUR students are found to be the least optimistic in their predictions regarding the treatment effects. The optimistic staff members of the ERCS and ZRCS are more accurate in their estimations compared to the WUR students. The forecast accuracy does not differ significantly between the NLRC staff members and the staff members of ERCS and ZRCS. Additionally, levels of familiarity with the context of this ICS intervention are found to be positively associated with forecast accuracy among NLRC staff members and staff members of other national societies.

Furthermore, the predicted effect size as well as the accuracy of predictions are found to be positively associated with reported confidence levels. It appears that respondents who are less confident about their predictions have the tendency to underestimate the impact of the specified treatment. The predictions regarding the effect of the specified treatment on child health are found to be an exception to this phenomenon. A lot of respondents have overestimated the impact of the treatment on child health. At the same time, these respondents indicated to be (relatively) confident about these predictions. Due to this error of judgment, the correlations between reported confidence levels and the accuracy of predictions regarding child health are found to be negative. However, these coefficients are relatively small and not statistically significant.

The results of the analysis of experts' beliefs confirm the difficulty of predicting impact and thereby demonstrate the importance of this impact study. An evaluation like this will inform the RC on the actual impact of their work. The outcomes of this study can be used to demonstrate impact to existing or potential donors and to inform effective strategies for maximising adoption rates of ICS.

However, since the external validity of this trial is assessed to be rather limited due to a number of reasons, we should be careful with generalizing the results of this study. For example, the most vulnerable households with under-fives are excluded from trial participation. As a result, the findings of this study on the effects of the randomly assigned treatments do not apply to this subset of households. Furthermore, the project villages have been selected based on the outcomes of an extensive vulnerability and resilience analysis. Subsequently, a wide range of awareness-raising activities have been implemented in these project villages during the last two years. This ICS intervention will thus be implemented in a highly specific context. Therefore, we cannot assume the treatment conditions to have similar effects when they will be implemented in isolation in other Ethiopian rural villages. Consequently, additional intervention-based studies in other populations need to be conducted to verify the results of this study.

The largest threat to the internal validity of this study is the introduction of response bias. Although precautionary measures are put in place, there is still a chance that respondents link the endline survey to the RC. If they do, their responses can be influenced by their dependencies, associations or former experiences with the RC. Furthermore, responses are potentially influenced by social desirability bias or feelings of resentment or demoralisation when subjects learn that they have been excluded from full treatment. However, the internal validity of this study is assessed to be relatively high, since the potential sources of bias are not expected to substantially affect the results of this study.

This impact study is designed in collaboration with the Ethiopian and Netherlands Red Cross Societies. It was quite a challenge to design the study in such a way that it is fully aligned with the way of working that is preferred by the involved RC Societies. The planning was an additional complicating factor. The implementation of the 3FM pneumonia project was already in full swing when the WUR was approached to design an impact study. During the process of designing this study, a continuous expansion of project activities took place. As a result, designing this study really has been a race against time. Besides the time constraints, we also had to take into account the budget constraints in the study design process.

It is now up to the ERCS to implement the ICS intervention. They will be supported in this process by the WUR and the NLRC. One year after the implementation of the ICS intervention, we will conduct the endline survey. The outcomes of this study will be used to design the ICS intervention that will be implemented in the sub-districts to which the 3FM pneumonia project has expanded throughout that year. Lessons learned can thus immediately be put into practice.

References

Alemayehu, Z.T., Seboka, Y., Tareqgne, A., Mekuanent, N., & Debalke, R. (2012). A report on controlled cooking test results made on multi-fuel and chimney integrated Injera baking stoves.

Amare, D., Endeblhatu, A., & Muhabaw, A. (2015). Enhancing biomass energy efficiency in rural households of Ethiopia. *Journal of Energy and Natural Resources*, *4*(2), 27-33.

Arif, T., Ashraf, A., Miller, G., Mobarak, A. M., Akter, N., Ali, A. M., ... & Dwivedi, P. (2011). Promotion of improved cookstove in rural Bangladesh. *Dhaka: BRAC. 30p (RED working paper no. 22)*.

Armstrong, J. R., & Campbell, H. (1991). Indoor AirPollution Exposure and Lower Respiratory Infections in Young Gambian Children. *International journal of epidemiology*, *20*(2), 424-429.

Ashraf, N., Berry, J., & Shapiro, J. M. (2010). Can higher prices stimulate product use? Evidence from a field experiment in Zambia. *American Economic Review*, *100*(5), 2383-2413.

Bang, A. T., Bang, R. A., Tale, O., Sontakke, P., Solanki, J., Wargantiwar, R., & Kelzarkar, P. (1990). Reduction in pneumonia mortality and total childhood mortality by means of community-based intervention trial in Gadchiroli, India. *The Lancet*, *336*(8709), 201-206.

Bailis, R., Drigo, R., Ghilardi, A., & Masera, O. (2015). The carbon footprint of traditional woodfuels. *Nature Climate Change*, *5*(3), 266-272.

Baqui, A. H., Zaman, K., Persson, L. A., Arifeen, S. E., Yunus, M., Begum, N., & Black, R. E. (2003). Simultaneous weekly supplementation of iron and zinc is associated with lower morbidity due to diarrhea and acute lower respiratory infection in Bangladeshi infants. *The Journal of nutrition*, *133*(12), 4150-4157. .

Barnes, B., Mathee, A., & Moiloa, K. (2005). Assessing child time–activity patterns in relation to indoor cooking fires in developing countries: a methodological comparison. *International Journal of Hygiene and Environmental Health*, *208*(3), 219-225.

Barnes, B., Mathee, A., & Thomas, E. (2011). The impact of health behaviour change intervention on indoor air pollution indicators in the rural North West Province, South Africa. *Journal of Energy in Southern Africa*, *22*(3), 35-44.

Barnes, B. R. (2014). Behavioural change, indoor air pollution and child respiratory health in developing countries: A review. *International journal of environmental research and public health*, *11*(5), 4607-4618.

Bensch, G., Grimm, M., & Peters, J. (2015). Why do households forego high returns from technology adoption? Evidence from improved cooking stoves in Burkina Faso. *Journal of Economic Behavior & Organization*, *116*, 187-205.

Bensch, G., & Peters, J. (2015). The intensive margin of technology adoption–Experimental evidence on improved cooking stoves in rural Senegal. *Journal of health economics*, *42*, 44-63.

Bensch, G., & Peters, J. (2017). One-off subsidies and long-run adoption–Experimental evidence on improved cooking stoves in Senegal.

Bizzarri, M. (2010). Safe access to firewood and alternative energy in Ethiopia: An appraisal report. *Prepared for the World Food Program.* Retrieved from http://www.genderconsult.org/uploads/publications/doc/SAFE_Ethiopia_Appraisa l_Report_Final_Draft_2. pdf.

Black, R.E., Brown, K., Becker, S. (1982). Longitudinal studies of infectious diseases and physical growth of children in rural Bangladesh. *Am J Epidemiol, 115*, 305–314.

Bluffstone, R. A., Beyene, A. D., Gebreegziabher, Z., Martinsson, P., Mekonnen, A., & Vieider, F. M. (2017). If people pay for improved biomass stoves, do they use them more frequently? Evidence from a field experiment in Ethiopia.

Bluffstone, R., Beyene, A., Gebreegziabher, Z., & Mekonnen, A. (2015). Learning, experience, fuelwood and time savings from improved biomass cookstoves: Evidence from randomized controlled cooking tests in Ethiopia. In *Paper Presented at the 21st Annual Conference of the European Association of Environmental and Resource Economists*.

Broor, S., Pandey, R. M., Ghosh, M., Maitreyi, R. S., Lodha, R., Singhal, T., & Kabra, S. K. (2001). Risk factors for severe acute lower respiratory tract infection in under-five children. *Indian pediatrics*, *38*(12), 1361-1369.

Bruce, N., Rehfuess, E., Mehta, S., Hutton, G., & Smith, K. (2006). Indoor air pollution.

Bruce, N., McCracken, J., Albalak, R., Scheid, M., Smith, K. R., Lopez, V., & West, C. (2004). Impact of improved stoves, house construction and child location on levels of indoor air pollution exposure in young Guatemalan children. *Journal of Exposure Science and Environmental Epidemiology*, *14*(S1), S26.

Cardoso, M. R. A., Cousens, S. N., de Góes Siqueira, L. F., Alves, F. M., & D'Angelo, L. A. V. (2004). Crowding: risk factor or protective factor for lower respiratory disease in young children?. *BMC Public Health*, *4*(1), 19.

Central Statistical Agency [Ethiopia] and ICF International. (2016). Ethiopia Demographic and Health Survey 2016. *Addis Ababa, Ethiopia, and Rockville, Maryland, USA: CSA and ICF.*

Central Statistical Agency [Ethiopia]. (2012). Ethiopia demographic and health survey 2011. *Addis Ababa, Ethiopia and Calverton, Maryland, USA: CSA and ICF International, 430.*

Chengappa, C., Edwards, R., Bajpai, R., Shields, K. N., & Smith, K. R. (2007). Impact of improved cookstoves on indoor air quality in the Bundelkhand region in India. *Energy for Sustainable Development*, *11*(2), 33-44.

Chowdhury, Z., Le, L. T., Masud, A. A., Chang, K. C., Alauddin, M., Hossain, M., ... & Hopke, P. K. (2012). Quantification of indoor air pollution from using cookstoves and estimation of its health effects on adult women in northwest Bangladesh. *Aerosol Air Qual Res*, *12*(4), 463-475.

Chowdhury, M. S. H., Koike, M., Akther, S., & Miah, D. (2011). Biomass fuel use, burning technique and reasons for the denial of improved cooking stoves by Forest User Groups of Rema-Kalenga Wildlife Sanctuary, Bangladesh. *International Journal of Sustainable Development & World Ecology*, *18*(1), 88-97.

Clark, M. L., Heiderscheidt, J. M., & Peel, J. L. (2015). Integrating behavior change theory and measures into health-based cookstove interventions: a proposed epidemiologic research agenda. *Journal of health communication*, *20*(sup1), 94-97.

Cohen, J., & Dupas, P. (2010). Free distribution or cost-sharing? Evidence from a randomized malaria prevention experiment. *The Quarterly Journal of Economics*, 1-45.

Coles, C. L., Fraser, D., Givon-Lavi, N., Greenberg, D., Gorodischer, R., Bar-Ziv, J., & Dagan, R. (2005). Nutritional status and diarrheal illness as independent risk factors for alveolar pneumonia. *American journal of epidemiology*, *162*(10), 999-1007.

Conley, T. G., & Udry, C. R. (2010). Learning about a new technology: Pineapple in Ghana. *American economic review*, *100*(1), 35-69.

Dawit, W. (2009). Fuel efficient technology adoption in Ethiopia: evidence from improved “mirt” stove technology: a case in selected kebeles from “Adea” wereda. *Ethiopian Journal of Economics*, *17*(2), 107-107.

Dresen, E., DeVries, B., Herold, M., Verchot, L., & Müller, R. (2014). Fuelwood savings and carbon emission reductions by the use of improved cooking stoves in an Afromontane Forest, Ethiopia. *Land*, *3*(3), 1137-1157.

Dupas, P. (2009). What matters (and what does not) in households' decision to invest in malaria prevention?. *American Economic Review*, *99*(2), 224-30.

Dupas, P., & Robinson, J. (2013). Why don't the poor save more? Evidence from health savings experiments. *American Economic Review*, *103*(4), 1138-71.

Deutsche Gesellschaft für Internationale Zusammenarbeit. (2011). Mirt Stove Ethiopia—Stove Testing Results. Retrieved from http://cdm.unfccc.int/filestorage/_/p/LRDOC4GHWKQI80BUYVZA2J96ETF513/Additional%20information%3A%20Mirt%20Stove.pdf?t=NlZ8bnoyNzllfDAYSlDC5dM6_B7 4OWk4OQ0A

Deutsche Gesellschaft für Internationale Zusammenarbeit (2013). Clean and efficient cooking energy for 100 million homes: Results from the Bonn International Cooking Energy Forum 26–28 June, 2013.

Dutta, K., Shields, K. N., Edwards, R., & Smith, K. R. (2007). Impact of improved biomass cookstoves on indoor air quality near Pune, India. *Energy for Sustainable Development*, *11*(2), 19-32.

Ethiopian Red Cross Society, & Netherlands Red Cross Society. (2017). [3FM pneumonia project baseline dataset]. Unpublished raw data.

Ezzati, M., & Kammen, D. M. (2001). Indoor air pollution from biomass combustion and acute respiratory infections in Kenya: an exposure-response study. *The Lancet*, *358*(9282), 619-624.

Fatmi, Z., & White, F. (2002). A comparison of 'cough and cold' and pneumonia: risk factors for pneumonia in children under 5 years revisited. *International Journal of Infectious Diseases*, *6*(4), 294-301.

Fekadu, G. A., Terefe, M. W., & Alemie, G. A. (2014). Prevalence of pneumonia among under-five children in Este Town and the surrounding rural Kebeles, Northwest Ethiopia: a community based cross sectional study. *Science Journal of Public Health*, *2*(3), 150-155.

Forests news. (2014). *Six pieces of concrete: How the Mirt stove works* [Picture]. Retrieved from https://forestsnews.cifor.org/25166/improved-cooking-stoves-ethiopia-kafa-forests

Fullerton, D. G., Bruce, N., & Gordon, S. B. (2008). Indoor air pollution from biomass fuel smoke is a major health concern in the developing world. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, *102*(9), 843-851.

Gaia Assocation. (2014). The holistic feasibility study of national scale-up program for ethanol cookstoves and Ethanol Micro Distilleries (EMDs) in Ethiopia marketing strategy for ethanol to household cooking fuel.

Giné, X., & Yang, D. (2009). Insurance, credit, and technology adoption: Field experimental evidence from Malawi. *Journal of development Economics*, *89*(1), 1-11.

Guiteras, R., Levinsohn, J., & Mobarak, A. M. (2015). Encouraging sanitation investment in the developing world: a cluster-randomized trial. *Science*, *348*(6237), 903-906.

Geary, C. W., Prabawanti, C., & Aristani, C. (2014). A field assessment of adoption of improved cookstove practices in Yogyakarta, Indonesia: Focus on structural drivers. *Fhi360.* Available: http://www.fhi360.org/sites/default/files/media/documents/CookstovePracticeFieldAssessment101512

Gebreegziabher, Z., Beyene, A. D., Bluffstone, R., Martinsson, P., Mekonnen, A., & Toman, M. A. (2018). Fuel savings, cooking time and user satisfaction with improved biomass cookstoves: Evidence from controlled cooking tests in Ethiopia. *Resource and Energy Economics*, *52*, 173-185.

Geleta, D., Tessema, F., & Ewnetu, H. (2016). Determinants of Community Acquired Pneumonia among Children in Kersa District, Southwest Ethiopia: Facility Based Case Control Study. *J Pediatr Neonatal Care*, *5*(2), 00179.

Goetghebuer, T., Kwiatkowski, D., Thomson, A., & Hull, J. (2004). Familial susceptibility to severe respiratory infection in early life. *Pediatric pulmonology*, *38*(4), 321-328.

Grant, C. C., Emery, D., Milne, T., Coster, G., Forrest, C. B., Wall, C. R., ... & Tukuitonga, C. (2012). Risk factors for community-acquired pneumonia in pre-school-aged children. *Journal of paediatrics and child health*, *48*(5), 402-412.

Hassan, M. A. K., & Al-Sadoon, I. (2001). Risk factors for severe pneumonia in children in Basrah. *Tropical doctor*, *31*(3), 139-141.

Hueso, A., & Bell, B. (2013). An untold story of policy failure: the Total Sanitation Campaign in India. *Water Policy*, *15*(6), 1001-1017.

Jackson, S., Mathews, K. H., Pulanić, D., Falconer, R., Rudan, I., Campbell, H., & Nair, H. (2013). Risk factors for severe acute lower respiratory infections in children–a systematic review and meta-analysis. *Croatian medical journal*, *54*(2), 110-121.

Jamali, T., Fatmi, Z., Shahid, A., Khoso, A., Kadir, M. M., & Sathiakumar, N. (2017). Evaluation of short-term health effects among rural women and reduction in household air pollution due to improved cooking stoves: quasi experimental study. *Air Quality, Atmosphere & Health*, *10*(7), 809-819.

Jetter, J., Zhao, Y., Smith, K. R., Khan, B., Yelverton, T., DeCarlo, P., & Hays, M. D. (2012). Pollutant emissions and energy efficiency under controlled conditions for household biomass cookstoves and implications for metrics useful in setting international test standards. *Environmental science & technology*, *46*(19), 10827-10834.

Johnson, M. A., & Chiang, R. A. (2015). Quantitative stove use and ventilation guidance for behavior change strategies. *Journal of health communication*, *20*, 6-9.

Kar, K., & Pasteur, K. (2005). Subsidy or self-respect?: Community led total sanitation; an update on recent developments.

Kőszegi, B., & Rabin, M. (2006). A model of reference-dependent preferences. *The Quarterly Journal of Economics*, *121*(4), 1133-1165.

Koyanagi, A., Humphrey, J. H., Ntozini, R., Nathoo, K., Moulton, L. H., Iliff, P., ... & ZVITAMBO Study Group. (2011). Morbidity among human immunodeficiency virus-exposed but uninfected, human immunodeficiency virus-infected, and human immunodeficiency virus-unexposed infants in Zimbabwe before availability of highly active antiretroviral therapy. *The Pediatric infectious disease journal*, *30*(1), 45-51.

Kremer, M., & Miguel, E. (2007). The illusion of sustainability. *The Quarterly journal of economics*, *122*(3), 1007-1065.

LaFave, D., Beyene, A. D., Bluffstone, R., TM, S., Dissanayake, Z. G., Mekonnen, A., & Toman, M. (2018). Impacts of Improved Biomass Cookstoves on Child and Adult Health: Experimental Evidence from Rural Ethiopia. *Waterville, ME*.

Leis, K. S., McNally, J. D., Montgomery, M. R., Sankaran, K., Karunanayake, C., & Rosenberg, A. M. (2012). Vitamin D intake in young children with acute lower respiratory infection. *Translational pediatrics*, *1*(1), 6.

Lewis, J. J., & Pattanayak, S. K. (2012). Who adopts improved fuels and cookstoves? A systematic review. *Environmental health perspectives*, *120*(5), 637-645.

Macedo, S. E. C., Menezes, A. M. B., Albernaz, E., Post, P., & Knorst, M. (2007). Risk factors for acute respiratory disease hospitalization in children under one year of age. *Revista de saúde pública*, *41*(3), 351-358.

Masera, O., Edwards, R., Arnez, C. A., Berrueta, V., Johnson, M., Bracho, L. R., ... & Smith, K. R. (2007). Impact of Patsari improved cookstoves on indoor air quality in Michoacán, Mexico. *Energy for Sustainable Development*, *11*(2).

McCracken, J. P., Smith, K. R., Díaz, A., Mittleman, M. A., & Schwartz, J. (2007). Chimney stove intervention to reduce long-term wood smoke exposure lowers blood pressure among Guatemalan women. *Environmental health perspectives*, *115*(7), 996-1001.

Miller, G., & Mobarak, A. M. (2013). Learning about new technologies through opinion leaders and social networks: experimental evidence on non-traditional stoves in rural Bangladesh. *Networks*.

Miller, G., & Mobarak, A. M. (2014). Learning about new technologies through social networks: experimental evidence on nontraditional stoves in Bangladesh. *Marketing Science*, *34*(4), 480-499.

Mobarak, A. M., Dwivedi, P., Bailis, R., Hildemann, L., & Miller, G. (2012). Low demand for nontraditional cookstove technologies. *Proceedings of the National Academy of Sciences*, *109*(27), 10815-10820.

Mortimer, K., Ndamala, C. B., Naunje, A. W., Malava, J., Katundu, C., Weston, W., ... & Wang, D. (2017). A cleaner burning biomass-fuelled cookstove intervention to prevent pneumonia in children under 5 years old in rural Malawi (the Cooking and Pneumonia Study): a cluster randomised controlled trial. *The Lancet*, *389*(10065), 167-175.

Moya, J., Bearer, C. F., & Etzel, R. A. (2004). Children's behavior and physiology and how it affects exposure to environmental contaminants. *Pediatrics*, *113*(Supplement 3), 996-1006.

Muhe, L., Lulseged, S., Mason, K. E., & Simoes, E. A. (1997). Case-control study of the role of nutritional rickets in the risk of developing pneumonia in Ethiopian children. *The Lancet*, *349*(9068), 1801-1804.

Mulholland, K. (1995). Measles and pertussis in developing countries with good vaccine coverage. *Lancet, 345*, 305–307.

Mulholland, K., Hilton, S., Adegbola, R., Usen, S., Oparaugo, A., Omosigho, C., ... & Lahai, G. (1997). Randomised trial of Haemophilus influenzae type-b tetanus protein conjugate for prevention of pneumonia and meningitis in Gambian infants. *The Lancet*, *349*(9060), 1191-1197.

Murray, E. L., Brondi, L., Kleinbaum, D., McGowan, J. E., Van Mels, C., Brooks, W. A., ... & Bridges, C. B. (2012). Cooking fuel type, household ventilation, and the risk of acute lower respiratory illness in urban Bangladeshi children: a longitudinal study. *Indoor air*, *22*(2), 132-139.

Nguyen, T. T. P. T. (2017). Women's adoption of improved cook stoves in Timor-Leste: challenges and opportunities. *Development in Practice*, *27*(8), 1126-1132.

Osendarp, S. J., Santosham, M., Black, R. E., Wahed, M. A., Van Raaij, J. M., & Fuchs, G. J. (2002). Effect of zinc supplementation between 1 and 6 mo of life on growth and morbidity of Bangladeshi infants in urban slums. *The American journal of clinical nutrition*, *76*(6), 1401-1408.

Oyejide, C.O., Osinusi, K. (1990). Acute respiratory tract infection in children in Idikan Community, Ibadan, Nigeria: severity, risk factors and severity of occurrence. *Rev Infect Dis, 12*, 1042–1046.

Parikh, J. (2011). Hardships and health impacts on women due to traditional cooking fuels: A case study of Himachal Pradesh, India. *Energy Policy*, *39*(12), 7587-7594.

Pattanayak, S. K., & Pfaff, A. (2009). Behavior, environment, and health in developing countries: evaluation and valuation. *Annu. Rev. Resour. Econ.*, *1*(1), 183-217.

Pine, K., Edwards, R., Masera, O., Schilmann, A., Marrón-Mares, A., & Riojas-Rodríguez, H. (2011). Adoption and use of improved biomass stoves in Rural Mexico. *Energy for sustainable development*, *15*(2), 176-183.

Rehfuess, E. A., Puzzolo, E., Stanistreet, D., Pope, D., & Bruce, N. G. (2013). Enablers and barriers to large-scale uptake of improved solid fuel stoves: a systematic review. *Environmental health perspectives*, *122*(2), 120-130.

Rudan, I., Boschi-Pinto, C., Biloglav, Z., Mulholland, K., & Campbell, H. (2008). Epidemiology and etiology of childhood pneumonia. *Bulletin of the world health organization*, *86*, 408-416B.

Rudan, I., O'brien, K. L., Nair, H., Liu, L., Theodoratou, E., Qazi, S., ... & Child Health Epidemiology Reference Group. (2013). Epidemiology and etiology of childhood pneumonia in 2010: estimates of incidence, severe morbidity, mortality, underlying risk factors and causative pathogens for 192 countries. *Journal of global health*, *3*(1) .

Savitha, M. R., Nandeeshwara, S. B., Kumar, M. P., & Raju, C. K. (2007). Modifiable risk factors for acute lower respiratory tract infections. *The Indian Journal of Pediatrics*, *74*(5), 477-482.

Schilmann, A., Riojas-Rodríguez, H., Ramírez-Sedeño, K., Berrueta, V. M., Pérez-Padilla, R., & Romieu, I. (2015). Children's respiratory health after an efficient biomass stove (Patsari) intervention. *EcoHealth*, *12*(1), 68-76.

Schmidt, W. P., Cairncross, S., Barreto, M. L., Clasen, T., & Genser, B. (2009). Recent diarrhoeal illness and risk of lower respiratory infections in children under the age of 5 years. *International journal of epidemiology*, *38*(3), 766-772.

Schlag, N., & Zuzarte, F. (2008). Market barriers to clean cooking fuels in sub-Saharan Africa: a review of literature. *Stockholm Environment Institute, Stockholm*.

Seguin, R., Flax, V. L., & Jagger, P. (2018). Barriers and facilitators to adoption and use of fuel pellets and improved cookstoves in urban Rwanda. *PloS one*, *13*(10), e0203775.

Selwyn, B.J. (1990). The epidemiology of acute respiratory tract infection in young children: Comparison of findings from several developing countries. *Rev Infect Dis, 12*, 870–888.

Shah, N., Ramankutty, V., Premila, P. G., & Sathy, N. (1994). Risk factors for severe pneumonia in children in south Kerala: a hospital-based case-control study. *Journal of tropical pediatrics*, *40*(4), 201-206.

Smith, K. R., McCracken, J. P., Weber, M. W., Hubbard, A., Jenny, A., Thompson, L. M., ... & Bruce, N. (2011). Effect of reduction in household air pollution on childhood pneumonia in Guatemala (RESPIRE): a randomised controlled trial. *The Lancet*, *378*(9804), 1717-1726.

Suzuki, M., Thiem, V. D., Yanai, H., Matsubayashi, T., Yoshida, L. M., Tho, L. H., ... & Ariyoshi, K. (2009). Association of environmental tobacco smoking exposure with an increased risk of hospital admissions for pneumonia in children under 5 years of age in Vietnam. *Thorax*, *64*(6), 484-489.

Tarozzi, A., Mahajan, A., Blackburn, B., Kopf, D., Krishnan, L., & Yoong, J. (2011). Micro-Loans, Insecticide-Treated Bednets and Malaria: Evidence from a Randomized Controlled Trial in Orissa (India). Working Paper.

Tugumisirize, F., Tumwine, J. K., & Mworoza, E. A. (2002). Missed opportunities and caretaker constraints to childhood vaccination in rural areas of Uganda. *East African medical journal*, *79*(7), 347-354.

Unicef. (2014, 12 November). *Ethiopia, 12 November 2014: Pneumonia is a leading single disease killing under-five children.* Retrieved from https://www.unicef.org/esaro/5440_eth2014_pneumonia.html

Verheggen, F. W., & van Wijmen, F. C. (1996). Informed consent in clinical trials. *Health Policy*, *36*(2), 131-153.

Victora, C. G., Fuchs, S. C., Flores, J. A. C., Fonseca, W., & Kirkwood, B. (1994). Risk factors for pneumonia among children in a Brazilian metropolitan area. *Pediatrics*, *93*(6), 977-985.

Walker, C. L. F., Rudan, I., Liu, L., Nair, H., Theodoratou, E., Bhutta, Z. A., ... & Black, R. E. (2013). Global burden of childhood pneumonia and diarrhoea. *The Lancet*, *381*(9875), 1405-1416.

Wayse, V., Yousafzai, A., Mogale, K., & Filteau, S. (2004). Association of subclinical vitamin D deficiency with severe acute lower respiratory infection in Indian children under 5 y. *European journal of clinical nutrition*, *58*(4), 563.

Wonodi, C. B., Deloria-Knoll, M., Feikin, D. R., DeLuca, A. N., Driscoll, A. J., Moïsi, J. C., ... & Scott, J. A. G. (2012). Evaluation of risk factors for severe pneumonia in children: the Pneumonia Etiology Research for Child Health study. *Clinical infectious diseases*, *54*(suppl_2), S124-S131.

Workeneh, G.J. (2005). Factors controlling households energy use: implication for the conservation of the environment. *Unpublished Msc Thesis, Addis Ababa University.*

World Food Programme, & Disaster Prevention and Preparedness Commission. (2003). *Ethiopia: Regions and Zones* [Map]. Retrieved from https://reliefweb.int/map/ethiopia/ethiopia-regions-and-zones

World Health Organization (WHO). (2016, September)**.** *Pneumonia.* Retrieved from http://www.who.int/mediacentre/factsheets/fs331/en/

World Health Organization. (2018, May 8). *Household air pollution and health.* Retrieved from https://www.who.int/news-room/fact-sheets/detail/household-air-pollution-and-health

Yosef, A. F. (2007). Assessing environmental benefits of Mirt stoves with particular reference to indoor air pollution (carbon monoxide and suspended particulate matter) and energy conservation. *Unpublished Msc thesis, Addis Ababa University*.

Appendix 1: Risk factors and indicators

Risk factor category	Risk factor variable	Source	Indicator from baseline dataset
Child demographics	Age	1	X
Birth details	Birth weight	2, 3, 4, 5, 9	X
Nutrition	Breastfeeding practices	2, 4, 6, 7, 18	Child has been breastfed or not
	Malnutrition (or undernutrition)	2, 3, 4, 5, 6, 8, 18, 19, 20	Child has showed symptoms of malnutrition during the last two weeks
	Zinc deficiency	6, 22, 23	X
	Vitamin D deficiency	20, 24, 25	X
Past morbidities or comorbidities	Diarrhoea	10	Child has showed symptoms of diarrhoea during the last two weeks
	Measles	6, 11	Child has showed symptoms of measles during the last two weeks
	AIDS/HIV infections	11, 12, 26	X
	Malaria	12	X
	History of pneumonia or wheezing	8, 5	X
Child's vaccination status	Measles	13	Child has received measles vaccine
	Pertussis	14	Child has received DPT-Hib-HepB vaccine
	Hib	15	Child has received DPT-Hib-HepB vaccine
	Child immunization coverage	3, 4, 7, 8, 27	The number of vaccination injections a child has had
Maternal characteristics	Maternal age	5, 16	Age of mother in years
	Maternal education	17, 28	Highest level of education completed by mother
	Knowledge on pneumonia prevention	-	Number of prevention measures mentioned by mother
Family environment	Crowding level	2, 3, 4, 5, 6, 8, 11, 29	Number of under-fives per room that is used for sleeping
	Parental smoking	11, 9, 30, 31	Regularity of smoking inside the house
	Exposure to indoor air pollution	4, 6, 11, 32, 33	Household cooks usually in the house and household uses mainly biomass for cooking.
	Upper respiratory infection in mother or siblings	7	X
Local health care system	Description of local health care system (quality of care, availability of treatments)	21	Description of care received for sick child from the health facility
	Access to health care (distance, time and costs to nearest HF)	21	Distance to health facility and total costs related to visit the HF for child (financial fee, travel and drugs)

Sources: [1]Shah, Ramankutty, Premila & Sathy (1994); [2]Rudan et al. (2013); [3]Coles et al. (2005); [4]Rudan, Boschi-Pinto, Biloglav, Mulholland & Campbell (2008); [5]Victora, Fuchs, Flores, Fonseca & Kirkwood (1994); [6]Goetghebuer, Kwiatkowski, Thomson & Hull (2004); [7]Walker et al. (2013); [8]Broor et al. (2001); [9]Grant et al. (2012); [10]Fatmi & White (2002); [11]Hassan & Al-Sadoon (2001); [12]Muhe, Lulseged, Mason & Simoes (1997); [13]Osendarp et al. (2002); [14]Baqui et al. (2003); [15]Wayse, Yousafzai, Mogale & Filteau (2004); [16]Leis et al. (2012); [17]Schmidt, Cairncross, Barreto, Clasen & Genser (2009); [18]World Health Organization (2016); [19]Black, Brown & Becker (1982); [20]Koyanagi et al. (2011); [21]Oyejide & Osinusi (1990); [22]Mulholland (1995); [23]Mulholland et al. (1997); [24] Savitha et al. (2007); [25]Selwyn (1990); [26]Wonodi et al. (2012); [27]Macedo et al. (2007); [28]Cardoso et al. (2004); [29]Armstrong & Campbell (1991); [30]Suzuki et al. (2009); [31]Murray et al. (2012); [32]Smith et al. (2011); [33]Bang et al. (1990)

Appendix 2: Descriptive statistics

Baseline data - General						
Level / Person	**Variable**	**n**	**Mean**	**St. Dev**	**Min**	**Max**
Household	Single-headed (*0=No, 1=Yes*)	452	0.30	0.46	0	1
	Nr. of under-fives	452	1.50	0.70	1	5
Respondent	Gender (*0=Male, 1=Female*)	452	0.87	0.34	0	1
Household head	Age in years	452	35.76	8.68	19	65
Caregiver	Age in years	452	28.47	7.05	16	75
Youngest child	Age in years	452	1.88	1.14	0	4

Baseline data - Household head and primary caregiver			
Variable	**Categories**	**n**	**%**
Household head identity		**452**	
	Father	377	83.41%
	Mother	70	15.49%
	Male relative	2	0.44%
	Female relative	2	0.44%
	Other Female	1	0.22%
Primary caregiver identity		**452**	
	Father	20	4.42%
	Mother	426	94.25%
	Male relative	1	0.22%
	Female relative	5	1.11%
Education level household head		**388**	
	Less than one year completed	136	35.05%
	Primary school	174	44.85%
	Secondary school	53	13.66%
	Higher education	25	6.44%
Education level primary caregiver		**346**	
	Less than one year completed	174	50.29%
	Primary school	122	35.26%
	Secondary school	37	10.69%
	Higher education	13	3.72%

Baseline data - Cooking fuel		
What type of fuel does your HH mainly use for cooking? (*n*= 452)	Freq	Percent
Wood	245	54.87%
Coal/Lignite	97	21.46%
Coal/Animal dung	7	1.55%
Animal dung	1	0.22%
Natural gas	16	3.54%
Electricity	1	0.22%
Straw/shrubs/grass	14	3.10%
Other	2	0.44%

Baseline data - Cooking location		
Where is the cooking usually done? (*n*= 452)	Freq	Percent
In a separated building	266	58.85%
In the house	124	27.43%
Outdoors	62	13.72%

Baseline data - Awareness on IAP		
Mentioned IAP reduction as pneumonia prevention measures? (*n*= 452)	Freq	Percent
No	387	85.62%
Yes	65	14.38%

Appendix 3: Characteristics of sub-districts

Zone	District	Sub-district	Prior projects	Presence of HF	Nr. under-fives	Nr. HH	Nr. sub-villages	Nr. HEWs	Nr. CBHFA volunteers	Distance to main road (in km)
South-Gondar	Simada	Sergawit	No	HP	922	1583	50	2	5	5
		Leul semira	No	HP	725	1245	38	2	5	24
		Gujign	No	HC and HP	993	1706	22	1	5	16
		Girariya	No	HP	541	930	39	1	5	14
		Arganashera	No	HP	767	1318	45	1	5	10
	Ebenat	Worgaja	PfR II, Neonatal project	HC and HP	606	1041	20	2	5	25
		Selamaya	No	HC and HP	1318	2264	39	2	5	25
		Jimman	No	HP	633	1087	22	1	5	5
		Gela matebiya	No	HC and HP	837	1438	19	2	5	22
		Wonberoch	PfR II, Neonatal project	HP	785	1347	26	2	5	7
Guji	Gorodolla	Mededun	WASH project	HP	226	168	6	4	10	55
		Chenmansa	WASH project	HP	182	165	6	4	10	35
		Belambel	WASH project	HP	123	758	29	2	10	30
		Gofehambo	WASH project	HP	201	84	3	5	0?	19
		Gennale	No	HC and HP	192	84	3	1	0?	47
	Girja	Kebenawa dima	WASH project	HC and HP	476	679	25	4	5	17
		Kelina Gobbu	No	HC and HP	461	381	14	3	5	12
		Gewisa	WASH project	HP	432	444	16	3	5	24
		Oda kelecha	WASH project	HP	435	319	11	3	5	17
		Gennale korecha	WASH project	HP	580	476	17		5	54