

BMJ Open

BMJ Open is committed to open peer review. As part of this commitment we make the peer review history of every article we publish publicly available.

When an article is published we post the peer reviewers' comments and the authors' responses online. We also post the versions of the paper that were used during peer review. These are the versions that the peer review comments apply to.

The versions of the paper that follow are the versions that were submitted during the peer review process. They are not the versions of record or the final published versions. They should not be cited or distributed as the published version of this manuscript.

BMJ Open is an open access journal and the full, final, typeset and author-corrected version of record of the manuscript is available on our site with no access controls, subscription charges or pay-per-view fees (<http://bmjopen.bmj.com>).

If you have any questions on BMJ Open's open peer review process please email info.bmjopen@bmj.com

BMJ Open

Does an innovative paper-based health information system (PHISICC) improve data quality and use in primary health care? Protocol of a multi-country, cluster randomised controlled trial in Sub-Saharan African rural settings.

Journal:	<i>BMJ Open</i>
Manuscript ID	bmjopen-2021-051823
Article Type:	Original research
Date Submitted by the Author:	02-Apr-2021
Complete List of Authors:	<p>Bosch-Capblanch, Xavier; Swiss Tropical and Public Health Institute Oyo-Ita, Angela; University of Calabar, Department of Community Medicine Muloliwa, Artur; Lurio University Faculty of Health Sciences Yapi, Richard; Centre Suisse de Recherches Scientifiques en Côte d'Ivoire Auer, Christian ; Swiss Tropical and Public Health Institute Samba, Mamadou; Ministère de la Santé et de l'Hygiène Publique Gajewski, Suzanne; Swiss Tropical and Public Health Institute Ross, Amanda; Swiss Tropical and Public Health Institute Krause, L Kendall; Bill & Melinda Gates Foundation Ekpenyong, Nnette; University of Calabar Teaching Hospital, Department of Community Medicine Nwankwo, Ogonna; University of Calabar, Department of Community Medicine Njepuome, Anthonia; Swiss Tropical and Public Health Institute Lee, Sofia; Swiss Tropical and Public Health Institute Sacarlal, Jahit ; Eduardo Mondlane University, Faculty of Medicine Madede, Tavares; Universidade Eduardo Mondlane Berté, Salimata; Centre Suisse de Recherches Scientifiques en Côte d'Ivoire Matsinhe, Graça; Ministério da Saúde Garba, Abdullahi; National Primary Healthcare Development Agency Brown, David; BCGI LLC / pivot-23.5°</p>
Keywords:	<p>Human resource management < HEALTH SERVICES ADMINISTRATION & MANAGEMENT, Organisation of health services < HEALTH SERVICES ADMINISTRATION & MANAGEMENT, Protocols & guidelines < HEALTH SERVICES ADMINISTRATION & MANAGEMENT, Quality in health care < HEALTH SERVICES ADMINISTRATION & MANAGEMENT, STATISTICS & RESEARCH METHODS</p>



1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

BMJ Open: first published as 10.1136/bmjopen-2021-051823 on 29 July 2021. Downloaded from <http://bmjopen.bmj.com/> on April 18, 2024 by guest. Protected by copyright.



I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our [licence](#).

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which [Creative Commons](#) licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

Title

Does an innovative paper-based health information system (PHISICC) improve data quality and use in primary health care? Protocol of a multi-country, cluster randomised controlled trial in Sub-Saharan African rural settings.

Authors

- Xavier Bosch-Capblanch: Corresponding author. Swiss Tropical and Public Health Institute. Basel, Switzerland. And University of Basel, Basel, Switzerland. x.bosch@unibas.ch
- Angela Oyo-Ita Department of Community Medicine, University of Calabar, Calabar, Nigeria; oyo_ita@yahoo.com
- Artur Manuel Muloliwa. Faculty of Health Sciences, Lúrio University, Nampula, Mozambique; muloliwa@yahoo.com.br
- Richard B Yapi. Centre Suisse de Recherches Scientifiques en Côte d'Ivoire, Abidjan, Côte d'Ivoire. And Centre d'Entomologie Médicale et Vétérinaire, Université Alassane Ouattara, Bouaké, Côte d'Ivoire. richard.yapi@csrs.ci
- Christian Auer. Swiss Tropical and Public Health Institute. Basel, Switzerland. And University of Basel, Basel, Switzerland. christian.auer@swisstph.ch
- Mamadou Samba. Ministère de la Santé et de l'Hygiène Publique, Abidjan, Côte d'Ivoire. And Université Félix Houphouët Boigny, Abidjan, Côte d'Ivoire. samba.mamadou@gmail.com
- Suzanne Gajewski. Swiss Tropical and Public Health Institute. Basel, Switzerland. And University of Basel, Basel, Switzerland. suzanne.gajewski@swisstph.ch.
- Amanda Ross. Swiss Tropical and Public Health Institute. Basel, Switzerland. And University of Basel, Basel, Switzerland. amanda.ross@swisstph.ch
- L Kendall Krause. Bill & Melinda Gates Foundation, Seattle, USA. Kendall.Krause@gatesfoundation.org
- Nnette Ekpenyong. Department of Community Medicine, University of Calabar Teaching Hospital, Calabar, Nigeria. And Department of Community Medicine University of Calabar, Calabar, Nigeria. nnekon2015@gmail.com
- Ogonna Nwankwo. Swiss Tropical and Public Health Institute, Basel, Switzerland. And University of Basel, Basel, Switzerland. And Health Policy, Management and Economics Firm, Department of Community. Medicine, University of Calabar, Calabar, Nigeria. ogonna.nwankwo@swisstph.ch
- Anthonia Ngozi Njebuome. Swiss Tropical and Public Health Institute, Abuja, Nigeria. ngonjep@yahoo.com.
- Sofia Mandjate Lee. Swiss Tropical and Public Health Institute; Manhiça, Mozambique. sofia_mandjate@yahoo.com.br.
- Jahit Sacarlal. Faculty of Medicine, Eduardo Mondlane University, Maputo, Mozambique. jahityash2002@gmail.com.
- Tavares Madede. Department of Community Health, Faculty of Medicine, Eduardo Mondlane University, Maputo, Mozambique. tmadede@gmail.com
- Salimata Berté. Centre Suisse de Recherches Scientifiques en Côte d'Ivoire, Abidjan, Côte d'Ivoire. sali.berte@csrs.ci
- Graça Matsinhe. Expanded Program on Immunization, Ministry of Health, Maputo, Mozambique. gmatsinhe@gmail.com.
- Abdullahi Bulama Garba. Director Planning, Research and Statistics, National Primary Health Care. Development Agency, Abuja, Nigeria. ab.garba@nphcda.gov.ng.
- David W Brown. Managing Principal, BCGI LLC / pivot-23.5°, North Carolina, USA. david.brown@pivot235.org

Abstract

Introduction

Frontline health workers in remote, rural health facilities are the first contact of the formal health sector and are confronted with life-saving clinical and public health decisions. Appropriate health information systems (HIS) support the collection and use of data, thus facilitating decision-making. However, HIS focus on reporting and are unfit to support critical decisions at the peripheral level. Since data tools are paper-based in most primary health care settings, we have produced an innovative paper-based HIS (PHISICC), embracing all health care areas in primary health care, using a Human Centred Design, co-creation approach. The PHISICC tools aid decision-making and include recording and reporting. We are carrying out a cluster-randomised controlled trial in three African countries to assess the effects of PHISICC compared with the current systems, on data use and quality, quality of care and health worker perceptions, in remote, rural settings.

Methods

We have selected study areas in rural zones of Côte d'Ivoire, Mozambique and Nigeria. Seventy health facilities in each country have been randomly allocated to using PHISICC tools or to continuing to use the regular HIS tools (35 per arm). We have selected three villages in the catchment area of each health facility to carry out surveys in 10 households each. Outcomes of interest include data quality and use, coverage of health services, health workers perceptions and other process and explanatory variables.

Discussion

We strive to contribute to producing robust evidence on health systems interventions, affecting people in remote, rural settings where the most vulnerable live. The PHISICC tools focus on decision-making rather than data and are meant to support health workers decisions as well as reporting to the higher levels of the system. Robust evidence on HIS can better find its way to high quality systematic reviews and guidance development to inform policy and practice.

Trial registration: Pan African Clinical Trials Registry - PACTR201904664660639. Registered 01/04/2019, <https://pactr.samrc.ac.za/Search.aspx>.

Keywords

Decision-making | Health Information Systems | Primary Health Care | Sub-Saharan Africa | Data quality | Quality of care | Human Centred Design

Article summary

Strengths and limitations of the study

- This research assesses the effects of paper-based health information systems, which are massively used particularly in remote, rural areas but which seriously neglected in research.
- The paper-based interventions have been developed using a human-centred design approach, with frontline health workers and designers driving the co-creation process.
- Despite the complexity of health systems interventions like this one, we have applied robust experimental methods, together with qualitative research, to assess and understand the effects of the paper-based intervention. Robust evidence on health systems is more likely to gain the credibility of policy-makers and to make it into systematic reviews, guidance development and policy and practice.
- Research targeting frontline health workers in remote, rural areas has to take place where they live and work, which poses serious obstacles in the organisation, management and monitoring of the trials.
- These obstacles, aggravated by the COVID-19 pandemic, have challenged the mobility of the research team, the availability of the intervention in one of the countries and the duration of the trials.

Introduction

Frontline health workers (HW) in remote, rural health facilities (HF) in many countries are the first contact with the formal health sector of the population and they are confronted with life-saving clinical and public health decisions on a daily basis. Decisions are made by exerting a balanced judgment on the information related to health care events, such as making the correct diagnoses or deciding on which vaccinations a child should receive on a given day. In order to properly handle this information, appropriate data support tools and processes are required, referred to as the health information system (HIS); or Routine HIS or Health Management Information System [1]. In reality, though, HIS are primarily designed to report aggregated health events to the higher tiers of the health systems rather than to inform decision-making at the point of care [2].

Increasing pressure by donors and governments to collect more and more data has aggravated the situation, through the proliferation of data support tools that have overloaded frontline health workers compromising their capacity to deliver good quality of care and to delivery good quality data [3], for higher level decision-making.

Promising 'quick fixes', such as the scale up of digital HIS, are taking a long time to implement and face enormous challenges related to infrastructure, equipment and services necessary to run them. Besides, research evidence on the effects of digital solutions remains patchy and inconsistent, even in high-income country settings, where complaints about computerisation of clinical care have been raised [4,5]. Hence, it is very likely that paper tools will remain a primary, if not unique, data support mechanism particularly in remote, rural HF in many countries.

PHISICC (Paper-based Health Information System in Comprehensive Care) is a multi-year, multi-country, mixed-methods research project that aims at producing and testing an innovative paper-based HIS to improve data quality and use, decision making and health outcomes, at Primary Health Care (PHC). It is being carried out in selected areas within Côte d'Ivoire, Mozambique and Nigeria. The project started in 2015, producing a systematic review on the effects of HIS interventions and a

1
2
3 framework synthesis on how HIS are understood in the literature. These were followed by studies to
4
5 characterise the existing HIS in the three countries. With these bodies of evidence, we engaged into a
6
7 Human Centred Design (HCD) co-creative process with frontline HW to design an innovative HIS
8
9 (PHISICC).
10

11
12 The impact of the PHISICC HIS on data quality and use, quality of health care and HW perceptions is
13
14 being assessed concurrently in rural areas in the three countries. We describe the design of the trial
15
16 here, consistent with CONSORT reporting guidelines [6] and the extension for cluster randomised
17
18 controlled trials (CRCT) [7]; see Additional file 1.
19
20

21 22 **Methods**

23 24 25 **Aim**

26
27 The aim of the trial is to address the research question: what are the effects of an innovative paper-
28
29 based HIS (PHISICC) on data use and quality, quality of health and HW perceptions compared with
30
31 the current HIS, in rural PHC settings?
32
33
34
35

36 37 **Patient and public involvement**

38
39 There was no public or patient involvement in this research because the intervention being assessed
40
41 in these trials target health care providers and decision-makers, rather than patients or the public in
42
43 general. We have involved health systems stakeholders and frontline health workers. Ministries of
44
45 Health at several levels participated in the preparation of the research proposal (personal
46
47 consultations), in the characterisation of health information systems that preceded the trials
48
49 (countries workshops), and throughout all project components (additional workshops, newsletters
50
51 and personal communication). Frontline health workers in the three countries have co-created the
52
53 intervention (i.e. paper based tools) through workshops, personal feedback and piloting under real
54
55 live conditions. Some of them are part of the research team and co-authoring this manuscript.
56
57
58
59
60

Study design

The study is a CRCT in each of the three countries. In each setting, 70 health facilities are randomised to intervention or control (35 per arm). The intervention arm HF use the new PHISICC tools (substituting the usual HIS tools) and the control arm HF use the regular HIS tools. The trial is implemented in the real life contexts of HF carrying out their usual duties.

The CRCT are implemented in the real life contexts of HF carrying out their usual duties. The trials started between the end of 2019 and beginning of 2020, depending on the country, when the intervention was installed and the baseline surveys carried out; and will last till mid-2021.

Study areas

Ministries of Health (MOH) officials in several countries were contacted before submitting the proposal to the funding agency in order to explore the willingness to engage in a project focusing on paper-based tools. Officials in several countries rejected the offer on the grounds of upcoming digitalisation plans of the HIS in the country. We partnered with MOH that found the research relevant to their context in three countries.

In each country, the eligibility criteria of study areas were that they had to belong to the operational area of research partners; contain a large enough number of health facilities and their catchment population; include vulnerable population (e.g. with low vaccination coverage, high childhood mortality); and be comparatively neglected in terms of infrastructure and services. We excluded areas with concurrent research or other types of activities that could conflict with the CRCT (such as the co-existence of another health-related study, massive developments in infrastructure or activities involving migration of the population, such as temporary work sites or changes in working sites) and areas with threats to safety or security that could jeopardise research activities.

The study areas are located in Adzopé, Agboville, Tiassalé and Sikensi districts (Côte d'Ivoire); in Funhalouro, Govuro, Homoine, Inhambane, Inharrime, Inhassoro, Mabote, Maxixe and Panda

1
2
3 (Inhambane province, Mozambique); and in Yala Local Government Authority (Cross-River State,
4
5 Nigeria).

6 7 8 **Eligibility of health facilities**

9
10 The intervention is implemented at the HF level. The eligibility criteria of the HF were that they had
11
12 to be located in the study areas, belong to the governmental health sector and their main activity
13
14 should be the delivery of PHC services. HF were excluded if they had specialised clinical services,
15
16 inpatients, physicians providing care or with plans for staff turn-over involving intervention and
17
18 control HF.
19

20
21 A 'master list' of eligible health facilities was prepared based on information provided by the MOH
22
23 across all study areas. We aimed at selecting 70 of the eligible HF in each country, using simple
24
25 random techniques in R [8].
26
27

28 29 30 **Allocation and blinding**

31
32 Allocation of the 70 HF per country into the intervention and control arms took place in a formal
33
34 event, gathering research partners and MOH officials to offer transparency and promote study
35
36 ownership by local and national authorities. Equally sized, folded pieces of paper with the names and
37
38 codes of included HF written on them were introduced in an opaque receptacle where they were
39
40 manually and blindly mixed. A second receptacle contained two equally sized pieces of paper, one
41
42 with the word 'intervention' and another one with the word 'control'. A selected person in the
43
44 meeting, not belonging to the research team, extracted one piece of paper at a time to reach half the
45
46 number of included HF. Then, a paper was extracted from the second receptacle to assign those HF
47
48 to the intervention or control arms. The rest of the papers were extracted as well to verify
49
50 completeness and no duplication of names, and those HF assigned to the other arm.
51
52

53
54 Once HF were selected, all villages or settlements for each health facility catchment area were listed
55
56 and three in each catchment area were selected. In practice, we selected all villages because the
57
58 numbers were below (in Côte d'Ivoire) or just above (in Nigeria) the needs. For each village, we used
59
60

1
2
3 Google satellite maps to identify and geo-locate every visible roof. Where there were many houses
4 per village (roughly, more than fifty or so), a researcher would mark four points in the map slightly
5 beyond the northernmost, southernmost, easternmost and westernmost roofs seen and 30 random
6 points were selected within that square. From the mapped points, 10 per village (with 10 more acting
7 as reserve) were randomly selected and marked on another map used in the field for data collectors
8 to approach households. Where technical problems impeded this approach in a given village, a field
9 supervisor would rotate a bottle on the floor towards the centre of the village and would select at
10 random 10 households in the direction pointed by the bottle, from the outer limit of the village till
11 the centre [9].
12
13
14
15
16
17
18
19
20
21
22

23 Blinding is only feasible for the research team members carrying out the CRCT data collection and the
24 analyses of the CRCT findings. The intervention (i.e. paper tools) are by design very different from the
25 existing system and it is not possible to blind participants or principal researchers.
26
27
28
29

30 We already had the agreement of the MOH and selected HF compliant with the inclusion criteria
31 were provided with the intervention shortly after completing the baseline data collection. Therefore,
32 recruitment as such took place at the same time of the allocation of HF into intervention and control
33 arms.
34
35
36
37
38
39

40 **The intervention**

41
42 The PHISICC paper-based intervention is a full set of paper-based tools to support decision-making by
43 frontline HW. These are the only tools to be used by HW in the intervention arm. The PHISICC tools
44 encompass the whole system (i.e. recording and reporting) and all clinical and public health care
45 areas and are characterised by: a common visual language (e.g. spaces for digits and text),
46 standardised formats across health care areas; support to critical data items (e.g. respiratory rate in
47 infants); graphic artefacts to distinguish severity degrees of signs or symptoms; documentation of
48 diagnoses and treatment decisions; and aides memoires, among others.
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 The PHISICC tools have been developed over 18 to 20 months prior to the CRCT, using a Human
4 Centred Design approach [10]. A strength of the Human Centred Design approach is its ability to
5 unlock the user's perspective so that designers can build solutions that are fully reality-based and
6 work well. Co-creation groups were formed in each country with researchers, staff from partner
7 institutions and healthcare workers, led by a team of professional designers. Based on co-creation,
8 participatory processes, and Human Centred Design principles, many iterations took place between
9 co-creation groups and end-users of the tools, the frontline HW, till reaching a design that
10 considered and addressed the main issues raised by HW (i.e. usability, clarity, size of tools). The
11 PHISICC tools have been produced in French for Côte d'Ivoire, in Portuguese for Mozambique and
12 English for Nigeria, which are the official languages used in the health systems in the three countries;
13 using the official logo of the MOHs. Health care areas covered include: family planning, antenatal
14 care, including tetanus toxoid vaccination, delivery, post-natal care, vaccination, sick child, adults
15 outpatient consultation, tuberculosis diagnosis and treatment, and HIV. Referral forms were also
16 designed.

17
18
19 The PHISICC tools have three sub-components: registers, tallies and reports. Registers are formed by
20 seven DIN-A3 and one DIN-A4 (for referrals) book covering all health care areas except for
21 tuberculosis treatment, for which DIN-A3 cards were used. Register books have 100, 200 or 400
22 pages depending on the country and health care area. They are used to record individual clients' data
23 for each health care event, either of clinical or public health nature. Some register books have clinical
24 notes at the very beginning, as 'aide memoires', and an example of a filled-in form, to assist HW
25 when doubting how to proceed.

26
27 Tallies are DIN-A3 single sheets which contain a list of the indicators to be transferred to higher levels
28 of the health system, with a series of small ovals, grouped in fives, to mark with tally sticks with a
29 pen. In contrast to the current systems that have no tallies or only for vaccination, tallies were
30 created for all health care areas. In the middle-right side of the tally, a column accommodates cells
31 aligned with the ovals to insert the count for each indicator; and in the far right of the sheet there is

1
2
3 a replica of the count column, separated with a perforated line, which is detached and sent, as part
4
5 of the monthly report to the higher level in the health system.
6

7
8 During three or four days, HW were trained on HIS before the start of the trial. In the intervention
9
10 arm they were trained on the PHISICC tools; and the control arm received a refresher training about
11
12 the regular tools, during the same number of days.
13

14
15 Additionally, given that the regular tools already contained information on past vaccination history of
16
17 children still to complete their vaccination schedule, we created a mechanism to retrieve data of
18
19 children's vaccination status to transcribe into the new vaccination register book in the intervention
20
21 arm ('system transition').
22

23
24 Tools were endorsed by MOH, printed in local printing companies and distributed to HW at the end
25
26 of the training sessions. A digital spreadsheet was created to monitor consumption and order
27
28 additional tools to cover health facility needs during the life of the trial.
29

30 31 **Outcomes**

32
33
34 There are five primary outcomes (Table 1). Vaccination adherence is defined as the total number of
35
36 vaccine doses given during the trial period in the correct time interval to children over the total
37
38 number of vaccine doses that should have been given during the same period. Antenatal care visits
39
40 uptake will also be considered depending on the expected number of pregnancies in the study areas.
41
42 Both are used as proxies for health outcomes in terms of protection against disease [11] and
43
44 prevention of pregnancy complications [12]. Data concordance is defined as the level of agreement
45
46 of HIS indicators between (i) records, (ii) tallies and (iii) reports [3]. In terms of data use for decision
47
48 making, we will estimate the diagnostics scope in the sick child (i.e. number of different diagnoses
49
50 per child; and treatment appropriateness (i.e. number of prescribed treatments that are supported
51
52 by a documented diagnosis). Health workers satisfaction will be assessed using a standardised
53
54 questionnaire [13,14,15]. While the intervention targets HF, some of the outcomes are measured at
55
56 the level of HF, and some from patients clustered within HF catchment areas.
57
58
59
60

1
2
3 Secondary outcomes are classified under the following domains: data quality, data user, mortality,
4
5 HW experience, clients experience and resource consumption:
6
7

- 8 • Data quality
 - 9
 - 10 ○ Completeness of recording and reporting in specific forms; i.e. prevalence of unduly
 - 11 missing data items; partograph used;
 - 12
 - 13 ○ accuracy of recorded figures in comparison to real events (e.g. physical counting of
 - 14 commodities, such as number of 500mg Paracetamol tablets as recorded versus
 - 15 number of 500mg Paracetamol tablets as counted;
 - 16
 - 17 ○ timeliness of reporting, as documented by time stamps in forms;
 - 18
 - 19 ○ loss of data or data which does not reach the next upper administrative level.
 - 20
 - 21
 - 22
 - 23
 - 24
 - 25
- 26 • Data use
 - 27
 - 28 ○ in terms of knowledge (e.g. vaccines due based on date of birth; weight for length
 - 29 assessments);
 - 30
 - 31 ○ cases of different conditions properly treated in (e.g. diarrhoea cases given oral
 - 32 rehydration therapy according to national guidelines; pneumonia cases given
 - 33 appropriate antibiotic according to national guidelines;
 - 34
 - 35 ○ public health decisions: availability of lost to follow up lists or plans for vaccination,
 - 36 tuberculosis and or HIV/AIDS treatment control;
 - 37
 - 38 ○ occurrence of stock outs of essential drugs
 - 39
 - 40
 - 41
 - 42
 - 43
 - 44
 - 45
- 46 • Overall under-5s mortality and under-5s mortality excluding peri-natal mortality [16].
- 47
- 48 • Health workers 'human experience' and satisfaction
- 49
- 50 • District health information officers' 'human experience'
- 51
- 52
- 53 • Clients 'human experience' and satisfaction
- 54
- 55
- 56
- 57
- 58
- 59
- 60

- Resources consumption (e.g. time use, costs)
 - intervention costs: tools, training, start-up;
 - time used for recording and reporting (e.g. time-motion study) [17];
 - cost-effectiveness per unit of additional improvement in outcomes of interest.

In addition, we will consider ‘explanatory outcomes’ that will help to understand how the measured effects have taken place and why. We will look at the details of the interplay between the intervention, the system, the users and the context. Process indicators will be based on the documented activities that have taken place, from the conception of the intervention, up to its implementation, monitoring and evaluation. Process indicators may include: intervention set up and implementation, monitoring of the use of the intervention, special activities targeted at vulnerable populations, district reactions related to the intervention, handling of data coming from the new system, sustainability based on costs information and perceptions, alignment with national health policies and donor priorities. We will also explore health care services characteristics looking at generic indicators from health facilities, such human resources profiles and relations with the communities, population characteristics and system and context characteristics captured in early stages of the project, where data are available.

Sample size calculations

The required sample sizes for each primary outcome were determined using simulation. This allowed us to account for levels of clustering (Table 1). We used the regression models detailed in the data section to analyse the simulated trials and estimate the power. The simulation code was written in R. For each country, we required the probability of a type I error (rejecting the null hypothesis when it is actually true) α to be less than 0.05 and a power of 80%.

The sampling frames are the study areas in each country, which include the health facilities and households in their catchment areas.

1
2
3 For vaccination adherence, using a sample size of 35 HF per arm, we would have 80% power in each
4 country to detect as significant a difference between a proportion of due vaccines given from 75% in
5 the control to 85% in the intervention arms, assuming one child per household, 30 households per HF
6 and a between-HF variation equivalent to a k of 0.1, where k is equal to the standard deviation
7 divided by the mean. The value of k is unknown, but was chosen in line with general observations by
8 Hayes and Bennet [18].
9

10
11 For data quality outcomes, with 35 HF per arm we would be able to detect as significant a difference
12 from a ratio of 0.7 (reported : recorded) vaccinations in the control arm to 0.8 with the intervention
13 with 80% power, assuming 100 recorded vaccinations per HF and a standard deviation of 0.1 in the
14 ratios between HF.
15

16
17 In terms of diagnostic scope, we would be able to detect an increase in the proportion of child-visits
18 with more than one diagnosis from 30% to 35% with 80% power with 35 HF per arm, 60 records per
19 HF and assuming a k of 0.1 [Error! Bookmark not defined.].
20

21
22 We would be able to detect as significant an increase from 50% of treatments having a
23 corresponding appropriate diagnosis to 60% with 80% power assuming 35 HF per arm, 1 treatment
24 per child, 25 children per HF and variation between HF corresponding to $k = 0.1$ [Error! Bookmark
25 not defined.].
26

27
28 For the outcome related to health workers' satisfaction, we would be able to detect as significant an
29 increase from 50% of health workers satisfied to 90%, with 80% power assuming 35 HF, three health
30 workers per HF and a variation between HF equivalent to $k = 0.1$.
31

32
33 In summary, in each country we require 35 HF per arm, three HW per HF, 100 vaccination records
34 per HF, 60 sick child records per health facility and 30 children per health facility catchment area.
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Data collection and management

Data collection took place at baseline and will take place again at the end of the study. Data is collected from health facilities, from the households in the catchment areas of the included health facilities and also from district offices.

For data quality and data use outcomes, HF registers, tallies and reports will be scrutinised. For population based outcomes, we carry out household surveys at baseline and at end-line. We use standard approaches for these types of surveys [19]. Households are visited, the research project is briefly introduced and consent requested.. Ideally, mothers of alive children or women in child-bearing age were interviewed in order to obtain information on living children (i.e. vaccination history) and death events, respectively, using home-based records if available and accessible. Patients' satisfaction will be assessed using the PSQ-18 satisfaction questionnaire [20,21,22]. Essentially, the tool enables practitioners to investigate the extent to which their health care service meets the perceived needs of their client group and pinpoint areas for improvement [22]. The interview will be conducted with consenting patients as close to their care encounter as possible [23]. Data tools are translated into the official languages of the study countries and pilot tested for consistent meaning and relevance to the setting. Data collectors are also able to communicate in local languages. The Satisfaction of Employees in Health Care (SEHC) survey is a validated tool to assess staff satisfaction. It was first developed and validated in a low-income country (Ethiopia) [24] and later successfully validated in a high-income country (USA) [25].

We use a mix of paper and electronic data (ODK [26]) collection tools. Data collectors are trained to minimise error. Tools are piloted before implementing. ODK data is regularly stored and sent to secure servers, as soon as data collectors reach their office base. Data from paper tools is double entered and compared and sent to secure servers. Each data collection tool has its corresponding electronic database that is cleaned and submitted to the analyses. All data is anonymised at the point of data collection or as soon as possible in the data management process. Data is labelled with an

1
2
3 arm code (e.g. 'A' or 'B') without any further information allowing to disclose which data items
4
5 belong to the intervention or to the control arms, ensuring blinding during data analyses.
6

7
8 Quality will be assured through several mechanisms: piloting of data collection tools; thorough
9
10 training of field workers; checking missing data related; double, independent data entry from papers
11
12 into digital databases; early descriptive analyses to detect potential outliers; fieldworkers tracking
13
14 and supervision.
15

16 17 18 **Data analysis**

19
20 The analysis will be carried out for each country separately, and based on intention-to-treat.

21
22 At baseline, data on population and health facility characteristics (i.e. basic demographic
23
24 characteristics of population and health workers, professional profile of health workers, health
25
26 facility size and services) will be produced and presented. If large imbalances are detected at
27
28 baseline, this information can be used to adjust the effect estimate comparisons [27,28].
29

30
31 The analyses vary for the different primary outcomes due to the unit of measurement and levels of
32
33 clustering, the type of variable, and whether measurements were taken at baseline and endpoint or
34
35 endpoint only. We use regression models to allow us to estimate the effect of the outcome while
36
37 flexibly accounting for these issues and allowing adjustment for potential confounders.
38

39
40 Logistic regression will be used for the binary variables: vaccine adherence is measured by
41
42 determining whether each vaccine due was received, and treatment appropriateness by whether
43
44 each treatment was correctly prescribed. Data concordance and diagnostic scope are count variables
45
46 and may be analysed with Poisson regression, depending on their distribution. The regression model
47
48 for HW satisfaction will depend on how it is distributed.
49

50
51 The outcomes have different levels of clustering (children or consultations, HW, HF). We will account
52
53 for these levels of clustering by including random effects in the regression models.
54

55
56 Four of the primary outcomes are measured at baseline and end-line. The effect of the intervention
57
58 will be estimated using an interaction term between arm and survey in the regression models: ie is
59
60

1
2
3 the change in the outcome between baseline and follow-up in the intervention arm different to the
4
5 change between baseline and follow-up in the control arm. The effect of HW satisfaction, measured
6
7 only at end-line, will be estimated as the difference between the intervention and control arm.
8
9

10 All estimates for the effect of the intervention will be presented with 95% confidence intervals. The
11
12 analyses will be carried out using R [29].
13
14

15 **Measures to minimise bias**

16
17 Statistical analyses will be carried out blindly, without knowledge of what health facilities or
18
19 population in the catchment area belong to the intervention or control groups. Only when the
20
21 analysis code is considered as definitive and fixed, will results be shared with the wider investigators
22
23 team and the arms for health facilities and population will be disclosed.
24
25

26
27 Outcome measurement bias may take place where data from the HIS, which is the focus of the
28
29 intervention, is used to measure outcomes. However, we will minimise this by assessing population
30
31 based outcomes at household level.
32
33

34 Contamination (i.e. the intervention affects individuals or units assigned to the control arm) may
35
36 happen via the exchanges between health workers from health facilities in both arms; for example:
37
38 in monthly district data quality meetings, managerial meetings; or through inputs from supervisors
39
40 who influence control health facilities with intervention tips encountered in health facilities of the
41
42 intervention arms. One mechanism to address this issue is using a district-based cluster
43
44 randomisation scheme. However, we consider that (i) contamination, despite increasing the
45
46 awareness of health works in control health facilities, will hardly influence the decision making
47
48 mechanisms that the HIS intervention focuses on; and (ii) randomisation at the level of district poses
49
50 additional challenges that are not worth the marginal benefit of reducing a doubtful contamination
51
52 [30].
53
54

55
56 The spill over effect (i.e. benefits of the intervention extend beyond their direct recipients) [31] may
57
58 take place in higher levels of the health systems; e.g. districts data managers and programme
59
60

1
2
3 managers may experience the benefits of better structured and more timely data produced in health
4 facilities in the intervention arms. The trial will have no capacity to quantitatively account for spill
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

managers may experience the benefits of better structured and more timely data produced in health facilities in the intervention arms. The trial will have no capacity to quantitatively account for spill overs at higher levels of the system, due to the limited number of higher level administrative areas that will be involved in the trial. However, through process indicators, we will consider potential benefits and harms of the intervention at higher levels of the system.

A challenge is the Hawthorne effect (i.e. observer effect). Both health workers in the intervention and in the control sites will have an awareness of being observed as data collection activities will be at the same level of intensity in the two arms. Therefore, there should be no differential effect.

Analyses will be based on the intention-to-treat. It is important to closely monitor if the intervention HFs consistently use the new HIS tools and approaches. The data collection team and the trial monitoring team will check if old forms are still being used in the intervention health facilities.

However, we do not expect health facilities to migrate between intervention and control arms, or vice versa, due to feasibility issues. On the other hand, some household members in a given catchment area may decide to seek for health care in a health facility belonging to another trial arm. In these cases, households will be analysed as belonging to the original trial arm.

Discussion

This is one of the very few studies assessing the effects of health systems interventions using experimental study designs [32]. Most of the experimental studies on HIS are circumscribed to specific health care areas (e.g. tuberculosis, vaccination, cardiovascular disease) and very few have a system-wide approach (e.g. PHC) [32]. This is the only experimental study we are aware of, focusing on paper-based HIS.

To date, some modifications to the protocol have taken place. In Côte d'Ivoire, we decided to select study areas close to the research institution base on logistics and practical reasons, instead of selecting an area in the north of the country, where poorer health indicators have been described. In Mozambique, the low density of HF per population implied extremely vast distances between HF and

1
2
3 this, coupled with the rainy season, made the trial unfeasible in the originally selected Nampula
4 province. After consultations, we decided to move the trial to the province of Inhambane and cancel
5 the household survey. The allocation of HF to the intervention and control arms was completed using
6 random number generation.
7
8
9

10
11
12 Experimental studies for health systems interventions are sometimes dismissed because of their
13 limited capacity to provide reliable explanations of complex health system issues [33]. While we
14 acknowledge these limitations, there is also a need for more robust evidence on the effects of these
15 types of health system interventions [34] and there are also good examples of experimental studies
16 reporting findings that can make it to the policy arena [35]. When embarking on this research, we
17 considered the type of evidence required to contribute to systematic reviews [36], guidance
18 development [37] and eventually recommendations for policy and practice [38]. Furthermore, we
19 have embedded the CRCT in a comprehensive research context which includes systematic reviews of
20 the literature and qualitative research, and we are also looking at explanatory outcomes within the
21 CRCT itself. We believe that this approach will provide a more comprehensive picture of what has
22 happened with the PHISICC tools used by HW and why.
23
24
25
26
27
28
29
30
31
32
33
34
35

36
37 We acknowledge the challenges of carrying out research on health care provided to remote, rural
38 communities (in this case in Sub-Saharan Africa). However, it is only in these remote areas where
39 research about their specific problems and needs can take place. Challenges included long distances,
40 poor conditions of roads, unreliable communications and limited food and accommodation services,
41 all of them to be proactively handled to keep the quality of work and the morale of researchers and
42 collaborators.
43
44
45
46
47
48
49

50
51 The engagement and ownership of partners within this research has also been instrumental in order
52 to plan and implement the CRCT. The intervention actually targets a governmental sub-system (the
53 HIS) for which we required more than permission but also endorsement and active support. We
54 achieved this level of collaboration by ensuring the participation of key stakeholders in key phases of
55
56
57
58
59
60

1
2
3 the whole project, from inception till the implementation of the last phases, through frequent
4
5 communication and workshops.
6

7
8 We are aware of the current trends towards digitalization of HIS. However, WHO recommendations
9
10 on the matter are weak since the underlying evidence to support these recommendations is
11
12 inconsistent [39]. The principles and methodological approaches in PHISICC can be applied to the
13
14 development of any technological solution, being on paper, digital or mixed. PHISICC, is not mainly
15
16 about technologies to support data, but rather about a change in paradigm where life-saving
17
18 decisions by frontline health workers are at the centre of the intervention; facilitating as well the
19
20 information requirements of higher levels in the health system.
21
22

23
24 We expect that the results of the trials, both quantitative and qualitative, will be able to inform
25
26 policies on how to make HIS responsive to providers' decision-making needs, particularly in health
27
28 services where the most vulnerable live.
29
30

31 **Authors' contributions**

32
33
34 XBC, AOI, AM, RBY and CAU prepared the proposal for the funding agency, conceived the study and
35
36 produced the data collection tools. SG ensured the regulatory, ethical and trial monitoring
37
38 components. AR developed the analytical approaches and made the sample size calculations. RBY,
39
40 MS and SB adapted the protocol to the context of Côte d'Ivoire and managed the administrative and
41
42 ethical approvals in the country; AOI, NE, OK, ANN and ABB likewise in Nigeria; AM, SML and GM, in
43
44 Nampula province (Mozambique); JS and TM adapted the protocol and acquired ethical and
45
46 administrative clearances for Inhambane province (Mozambique). DB is chair of the PHISICC
47
48 Technical Advisory Group (TAG) and has coordinated multiple formal and informal inputs. LKK and DB
49
50 have advised on the adequacy of the study protocol within the overall PHISICC proposal and TAG
51
52 advices. All country teams participated in PHISICC workshops and ensured that the protocol was
53
54 suitable to countries realities; developed data collection tools and training materials. They are
55
56
57
58
59
60

1
2
3 responsible for the implementation of the trial in each country. XBC drafted the first version of the
4
5 manuscript. All authors commented on several versions of the manuscript.
6
7

8 **Acknowledgements**

9
10 Members of the PHISICC Technical Advisory Group, chaired by David Brown: Blanche Anya (WHO
11
12 AFRO), Abdallah Bchir (former Gavi), Marta Gacic-Dobo (WHO HQ), Richard Greffé (AIGA), Pamela
13
14 Mitula (WHO AFRO), Sandy Oliver (UCL Institute) and Chris Wolf (BGMF).
15
16

17
18 Research collaborators: Momade Ali, Celso Belo, Bassirou Bonfoh, Lisa Diallo, John Ferreira, Bernard
19
20 Guessanbi, Caitlin Jarrett, Inza Koné, Felix Malé, Kouadio M'bra, David O'Donnell, Damaris Rodríguez
21
22 (Sonder Collective), Melanie Wendland and Meike Zuske (Swiss TPH).
23
24

25 Stakeholders from Côte d'Ivoire, Mozambique and Nigeria, participating in the consultation
26
27 processes.
28
29

30 **Data sharing statement**

31
32 We report on a protocol of three Cluster Randomised Controlled Trials and therefore no data is
33
34 available yet. On completion of the trials, access to data will be available from the national research
35
36 institutions in Côte d'Ivoire, Mozambique and Nigeria, in publications and in funding agency. Data
37
38 will be made available via an online data repository. Access will be granted following a review of
39
40 requests by SwissTPH contract officer.
41

42 **Ethical approvals in Côte d'Ivoire, Mozambique and Nigeria**

- 43 - Comité National Ethique des Sciences de la Vie et de la Santé (CNESVS), reference: 024-
44
45 19/MSHP/CNESVS-kp (Côte d'Ivoire)
- 46 - Comité Institucional de Bioética para Saúde da Universidade Lúrio, reference: 16.2/Julho/CBISUL/19
47
48 (Mozambique)
- 49 - Secretary, Government of Cross River State of Nigeria, Ministry of Health, Calabar Health Research
50
51 Ethics Committee, reference: CRS/MH/HREC/018/Vol. V1/151 (Nigeria)
- 52
53 - Ethikkommission Nordwest- und Zentralschweiz (EKNZ), reference: 2018-01059 (Switzerland).
54
55
56
57

58 **Funding**

59
60 Funder by the Bill & Melinda Gates Foundation, grant number INV-010193 / OPP1135947.

Abbreviations

BMGF Bill & Melinda Gates Foundation

CIV Côte d'Ivoire

CRCT Cluster Randomised Controlled Trial

sd Standard deviation

HF Health Facility

HW Health worker

MOZ Mozambique

NGA Nigeria

WHO World Health Organisation

Additional files

- Additional file 1: CONSORT statement checklist.

Competing interests

No, there are no competing interests for any author.

References

- 1 WHO. Monitoring results with health facility information systems: A technical consultation Glion-sur-Montreux, Switzerland, 11–12 June 2014. WHO 2014.
- 2 WHO. Health facility & community data toolkit. WHO 2014.
- 3 Bosch-Capblanch X, Ronveaux O, Doyle V, Remedios V, Bchir A. Accuracy and quality of immunization information systems in forty-one low income countries. *Trop Med Int Health*. 2009 Jan;14(1):2-10. doi: 10.1111/j.1365-3156.2008.02181.x. PMID: 19152556.
- 4 Loder E. Promises, promises. *BMJ*. 2017 Mar 23;356:j1446. doi: 10.1136/bmj.j1446. PMID: 28336513.
- 5 Wenger N, Méan M, Castioni J, Marques-Vidal P, Waeber G, Garnier A. Allocation of Internal Medicine Resident Time in a Swiss Hospital: A Time and Motion Study of Day and Evening Shifts. *Ann Intern Med*. 2017 Apr 18;166(8):579-586. doi: 10.7326/M16-2238. Epub 2017 Jan 31. PMID: 28135724.
- 6 Schulz KF, Altman DG, Moher D, for the CONSORT Group. CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials.
- 7 Campbell MK, Piaggio G, Elbourne DR, Altman DG; for the CONSORT Group. Consort 2010 statement: extension to cluster randomised trials. *BMJ*. 2012;345:e5661.
- 8 R Core Team. R: A Language and Environment for Statistical Computing [Internet]. Vienna, Austria; 2020. Available from: <https://www.R-project.org/>
- 9 Bostoen K, Chalabi Z. Optimization of household survey sampling without sample frames. *Int J Epidemiol*. 2006 Jun;35(3):751-5. doi: 10.1093/ije/dyl019. Epub 2006 Feb 15. PMID: 16481364.
- 10 Johnson, Sandhu, Tyler, “The Next Step for Human-Centered Design in Global Public Health” *Stanford Social Innovation Review*, Oct. 22, 2019.
- 11 WHO. Global Vaccine Action Plan 2011-2020. WHO 2013.
- 12 Carroli G, Villar J, Piaggio G, Khan-Neelofur D, Gülmezoglu M, Mugford M, Lumbiganon P, Farnot U, Bersgjø P; WHO Antenatal Care Trial Research Group. WHO systematic review of randomised controlled trials of routine antenatal care. *Lancet*. 2001 May 19;357(9268):1565-70. doi: 10.1016/S0140-6736(00)04723-1. PMID: 11377643.
- 13 Alpern R, Canavan ME, Thompson JT, McNatt Z, Tatak D, Lindfield T, Bradley EH. Development of a brief instrument for assessing healthcare employee satisfaction in a low-income setting. *PLoS One*.

1
2
3
4
5
6
7 2013 Nov 5;8(11):e79053. doi: 10.1371/journal.pone.0079053. PMID: 24223878; PMCID:

8
9 PMC3818514.

10 14Doll WJ and Torkzadeh G. The measurement of end-user computing satisfaction. *MIS Quarterly*
11 1988;12(2):259–74; 19.

12
13
14 15 Ilias A, Yasoa MR, Razaka MZA and Rahmana RA. The study of end-user computing satisfaction
16 (EUCS) on computerised accounting system (CAS) among Labuan F.T. Government sectors: a case
17 study in the responsibility centres. *Labuan e-Journal of Muamalat and Society* 2007;1:1–13.

18
19
20 16 Arifeen SE, Hoque DM, Akter T, Rahman M, Hoque ME, Begum K, Chowdhury EK, Khan R, Blum LS,
21 Ahmed S, Hossain MA, Siddik A, Begum N, Sadeq-ur Rahman Q, Haque TM, Billah SM, Islam M, Rumi
22 RA, Law E, Al-Helal ZA, Baqui AH, Schellenberg J, Adam T, Moulton LH, Habicht JP, Scherpbier RW,
23 Victora CG, Bryce J, Black RE. Effect of the Integrated Management of Childhood Illness strategy on
24 childhood mortality and nutrition in a rural area in Bangladesh: a cluster randomised trial. *Lancet*.
25 2009 Aug 1;374(9687):393-403. doi: 10.1016/S0140-6736(09)60828-X. PubMed PMID: 19647607.

26
27
28 17 Tilahun H, Fekadu B, Abdisa H, Canavan M, Linnander E, Bradley EH, Berman P. Ethiopia's health
29 extension workers use of work time on duty: time and motion study. *Health Policy Plan*. 2017 Apr
30 1;32(3):320-328. doi: 10.1093/heapol/czw129. PubMed PMID: 27658649.

31
32
33 18 Hayes RJ, Bennett S. Simple sample size calculation for cluster-randomized trials. *Int J Epidemiol*.
34 1999 Apr;28(2):319-26.

35
36
37 19 The DHS program. Demographic and Health surveys. <https://dhsprogram.com/>.

38
39
40 20 Thayaparan, A. a. (2013). The Patient Satisfaction Questionnaire Short Form (PSQ-18) as an
41 adaptable, reliable, and validated tool for use in various settings. *Med Educ Online*, 18.

42
43
44 21 The Health Foundation. (2013, June). Measuring patient experience (no.18). Retrieved Nov 6,
45 2018, from www.health.org.uk: [https://www.health.org.uk/publication/measuring-patient-](https://www.health.org.uk/publication/measuring-patient-experience)
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
experience.

22 Grogan, S. C. (2000). Validation of a questionnaire measuring patient satisfaction with general
practitioner services. *Quality in Health Care*, 9, 210-215.

23 LaVela, S. L. (2014). Evaluation and measurement of patient experience. *Patient Experience*
Journal, 1(1), Article 5.

24Alpern, R. C. (2013, November). Development of a Brief Instrument for Assessing Healthcare
Employee Satisfaction in a Low-Income Setting. *PLOS ONE*, 8(11), e79053.

- 1
2
3
4
5
6
7 25 Chang, E. C. (2017). Measuring job satisfaction among healthcare staff in the United States: a
8 confirmatory factor analysis of the Satisfaction of Employees in health Care (SEHC) survey.
9 International Journal for Quality in health Care, 29(2), 262-268.
10
11 26 Open Data Kit. <https://opendatakit.org/>. Accessed 5th February 2021.
12
13 27 Pocock S, et al. Subgroup analysis, covariate adjustment and baseline comparisons in clinical trial
14 reporting: current practice and problems. Stat Med. 2002;21:2917-30.
15
16 28 The CONSORT Group. CONSORT Statement [Internet]; item 15.
17
18 29 R Core Team. R: A language and environment for statistical computing. R Foundation for
19 Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.
20
21 30 Torgerson David J. Contamination in trials: is cluster randomisation the answer? BMJ 2001;
22 322:355.
23
24 31 Benjamin-Chung J, Arnold BF, Berger D, Luby SP, Miguel E, Colford JM Jr, Hubbard AE. Spillover
25 effects in epidemiology: parameters, study designs and methodological considerations. Int J
26 Epidemiol. 2018 Feb 1;47(1):332-347. doi: 10.1093/ije/dyx201. PubMed PMID: 29106568; PubMed
27 Central PMCID: PMC5837695.
28
29 32 Leon N, Balakrishna Y, Hohlfeld A, Odendaal WA, Schmidt BM, Zweigenthal V, Anstey Watkins J,
30 Daniels K. Routine Health Information System (RHIS) improvements for strengthened health system
31 management. Cochrane Database Syst Rev. 2020 Aug 13;8:CD012012. doi:
32 10.1002/14651858.CD012012.pub2. PMID: 32803893.
33
34 33 Allanson ER, Tunçalp Ö, Vogel JP, Khan DN, Oladapo OT, Long Q, Gülmezoglu AM. Implementation
35 of effective practices in health facilities: a systematic review of cluster randomised trials. BMJ Glob
36 Health. 2017 Jul 20;2(2):e000266. doi: 10.1136/bmjgh-2016-000266. PMID: 29081997; PMCID:
37 PMC5656132.
38
39 34 Brown DW, Bosch-Capblanch X, Shimp L. Where Do We Go From Here? Defining an Agenda for
40 Home-Based Records Research and Action Considering the 2018 WHO Guidelines. Glob Health Sci
41 Pract. 2019 Mar 29;7(1):6-11. doi: 10.9745/GHSP-D-18-00431. PMID: 30877139; PMCID:
42 PMC6538131.
43
44 35 Joos O, Silva R, Amouzou A, Moulton LH, Perin J, Bryce J, et al. Evaluation of a mhealth data
45 quality intervention to improve documentation of pregnancy outcomes by health surveillance
46 assistants in Malawi: a cluster randomized trial. PloS One 2016;11(1):e0145238. PMID: 26731401.
47
48 DOI: <https://doi.org/10.1371/journal.pone.0145238>
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3
4
5
6
7 36 Langlois EV, Ranson MK, Bärnighausen T, Bosch-Capblanch X, Daniels K, El-Jardali F, Ghaffar A,
8 Grimshaw J, Haines A, Lavis JN, Lewin S, Meng Q, Oliver S, Pantoja T, Straus S, Shemilt I, Tovey D,
9 Tugwell P, Waddington H, Wilson M, Yuan B, Røttingen JA. Advancing the field of health systems
10 research synthesis. *Syst Rev*. 2015 Jul 10;4:90. doi: 10.1186/s13643-015-0080-9. PMID: 26159806;
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

PMCID: PMC4498528.

37 Bosch-Capblanch X, Lavis JN, Lewin S, Atun R, Røttingen JA, Dröschel D, Beck L, Abalos E, El-Jardali
F, Gilson L, Oliver S, Wyss K, Tugwell P, Kulier R, Pang T, Haines A. Guidance for evidence-informed
policies about health systems: rationale for and challenges of guidance development. *PLoS Med*.
2012;9(3):e1001185. doi: 10.1371/journal.pmed.1001185. Epub 2012 Mar 6. PMID: 22412356;
PMCID: PMC3295823.

38 WHO. WHO recommendations on home-based records for maternal, newborn and child health.
WHO 2018.

39 WHO. Recommendations on digital interventions for health system strengthening. WHO
guideline. WHO 2019.

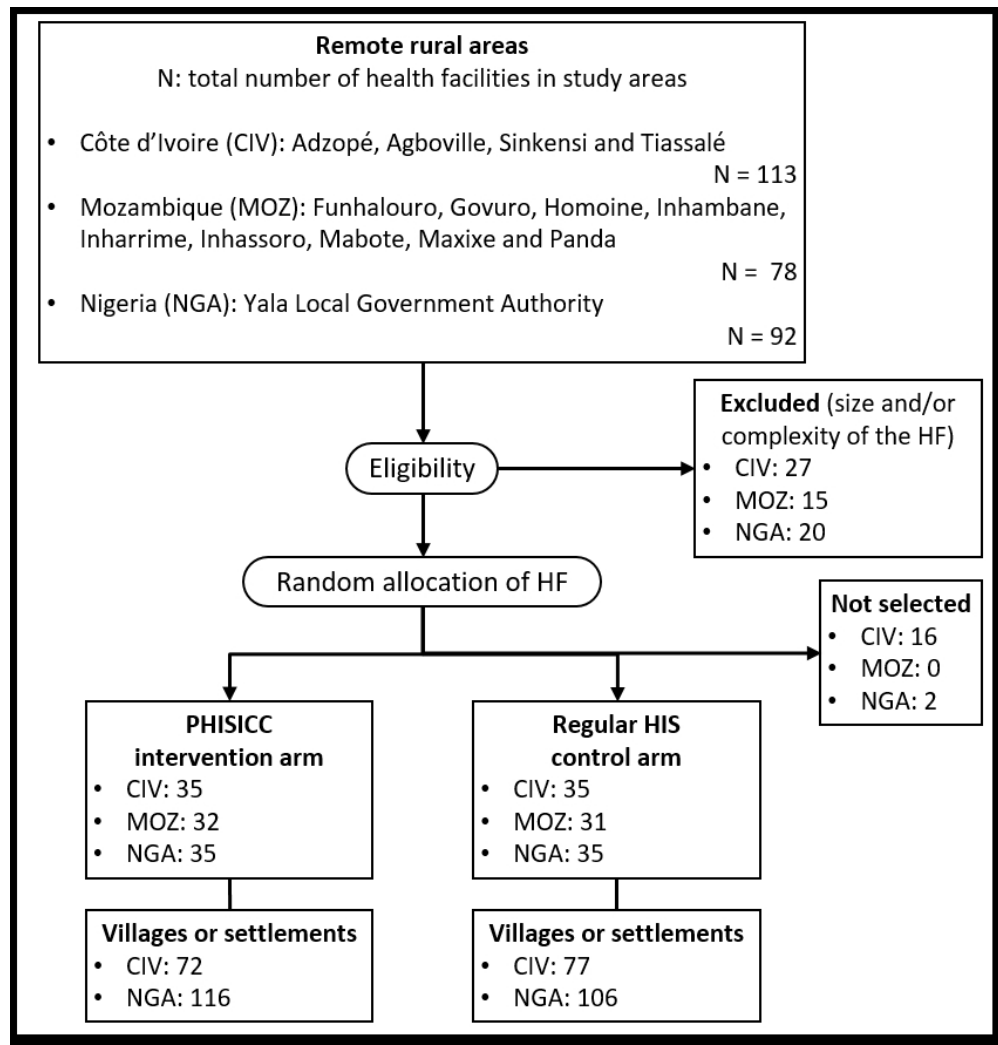
Tables and Figures

Table 1. Outcomes and parameters used to estimate the sample sizes.

	Outcome name	Subjects	Definition	Baseline estimate	Expected change	Comments
1	Vaccination adherence	Children under-1 in the households	Number of vaccines given in the previous calendar year over the number of vaccine due in the same period	75 given per 100 due	Increase of 10 per 100	Vaccines are clustered within children, and children within HFs
2	Data concordance	Recording tools in health facilities	Number of health care events (e.g. vaccinations, antenatal care consultations) recounted in the previous calendar year versus the number of health care events reported in the same time period	7 recounted for each 10 reported [3]	Increase of 2 recounted	A single estimate can be obtained in each HF or by time periods (no clustering)
3	Diagnostic scope	Records of sick child consultations	Number of diagnosis in each sick child consultation during the previous calendar year	1 or 2 per child	30% to 35% with more than 1 diagnosis	Individual consultations are clustered within HF
4	Treatment appropriateness	Records of sick child consultations	Number of treatments correctly prescribed in each sick child consultation during the previous calendar year	Half appropriate over all consultations	Increase to three quarters appropriateness	Individual consultations are clustered within HF
5	Health workforce satisfaction	Health workers	Degree (score) of satisfaction across all health facilities in each arm, with the intervention	5 out of 10	9 out of 10	Maybe two or three health workers can be approached in each health facility

Figure 1. CONSORT diagram: trial flow chart.

Separate file.



144x150mm (150 x 150 DPI)

Table 1: CONSORT 2010 checklist of information to include when reporting a cluster randomised trial

Section/Topic	Item No	Standard Checklist item	Extension for cluster designs	Page No *
Title and abstract				
	1a	Identification as a randomised trial in the title	Identification as a cluster randomised trial in the title	1
	1b	Structured summary of trial design, methods, results, and conclusions	See table 2	5
Introduction				
Background and objectives	2a	Scientific background and explanation of rationale	Rationale for using a cluster design	4
	2b	Specific objectives or hypotheses	Whether objectives pertain to the cluster level, the individual participant level or both	5
Methods				
Trial design	3a	Description of trial design (such as parallel, factorial) including allocation ratio	Definition of cluster and description of how the design features apply to the clusters	6
	3b	Important changes to methods after trial commencement (such as eligibility criteria), with reasons		Not applicable
Participants	4a	Eligibility criteria for participants	Eligibility criteria for clusters	6
	4b	Settings and locations where the data were collected		6
Interventions	5	The interventions for each group with sufficient details to allow replication, including how and when they were actually administered	Whether interventions pertain to the cluster level, the individual participant level or both	8
Outcomes	6a	Completely defined pre-specified primary and secondary outcome measures, including how and when they were assessed	Whether outcome measures pertain to the cluster level, the individual participant level or both	10 and Table 1
	6b	Any changes to trial outcomes after the trial commenced, with reasons		Not applicable
Sample size	7a	How sample size was determined	Method of calculation, number of clusters(s) (and whether equal or unequal cluster sizes are assumed), cluster size, a coefficient of intracluster correlation (ICC or k), and an indication of its uncertainty	12
	7b	When applicable, explanation of any interim analyses and stopping guidelines		Not yet applicable

Section/Topic	Item No	Standard Checklist item	Extension for cluster designs	Page No *
Randomisation:				
Sequence generation	8a	Method used to generate the random allocation sequence		7, 8
	8b	Type of randomisation; details of any restriction (such as blocking and block size)	Details of stratification or matching if used	Not applicable
Allocation concealment mechanism	9	Mechanism used to implement the random allocation sequence (such as sequentially numbered containers), describing any steps taken to conceal the sequence until interventions were assigned	Specification that allocation was based on clusters rather than individuals and whether allocation concealment (if any) was at the cluster level, the individual participant level or both	7
Implementation	10	Who generated the random allocation sequence, who enrolled participants, and who assigned participants to interventions	Replace by 10a, 10b and 10c	7
	10a		Who generated the random allocation sequence, who enrolled clusters, and who assigned clusters to interventions	7
	10b		Mechanism by which individual participants were included in clusters for the purposes of the trial (such as complete enumeration, random sampling)	10
	10c		From whom consent was sought (representatives of the cluster, or individual cluster members, or both), and whether consent was sought before or after randomisation	8
Blinding	11a	If done, who was blinded after assignment to interventions (for example, participants, care providers, those assessing outcomes) and how		8
	11b	If relevant, description of the similarity of interventions		Not applicable
Statistical methods	12a	Statistical methods used to compare groups for primary and secondary outcomes	How clustering was taken into account	15
	12b	Methods for additional analyses, such as subgroup analyses and adjusted analyses		15

Section/Topic	Item No	Standard Checklist item	Extension for cluster designs	Page No *
Results				Not yet applicable (protocol manuscript)
Discussion				
Limitations	20	Trial limitations, addressing sources of potential bias, imprecision, and, if relevant, multiplicity of analyses		18
Generalisability	21	Generalisability (external validity, applicability) of the trial findings	Generalisability to clusters and/or individual participants (as relevant)	18,19
Interpretation	22	Interpretation consistent with results, balancing benefits and harms, and considering other relevant evidence		Not yet applicable
Other information				
Registration	23	Registration number and name of trial registry		1
Protocol	24	Where the full trial protocol can be accessed, if available		1
Funding	25	Sources of funding and other support (such as supply of drugs), role of funders		20

* Page numbers: as seen in the document "draft_Proof_hi.pdf" (which has 34 pages)

BMJ Open

Does an innovative paper-based health information system (PHISICC) improve data quality and use in primary health care? Protocol of a multi-country, cluster randomised controlled trial in Sub-Saharan African rural settings.

Journal:	<i>BMJ Open</i>
Manuscript ID	bmjopen-2021-051823.R1
Article Type:	Protocol
Date Submitted by the Author:	02-Jun-2021
Complete List of Authors:	<p>Bosch-Capblanch, Xavier; Swiss Tropical and Public Health Institute; University of Basel Oyo-Ita, Angela; University of Calabar, Department of Community Medicine Muloliwa, Artur; Lurio University Faculty of Health Sciences Yapi, Richard; Centre Suisse de Recherches Scientifiques en Côte d'Ivoire Auer, Christian ; Swiss Tropical and Public Health Institute; University of Basel Samba, Mamadou; Ministère de la Santé et de l'Hygiène Publique Gajewski, Suzanne; Swiss Tropical and Public Health Institute; University of Basel Ross, Amanda; Swiss Tropical and Public Health Institute; University of Basel Krause, L Kendall; Bill & Melinda Gates Foundation Ekpenyong, Nnette; University of Calabar Teaching Hospital, Department of Community Medicine Nwankwo, Ogonna; University of Calabar, Department of Community Medicine Njepuome, Anthonia; Swiss Tropical and Public Health Institute Lee, Sofia; Swiss Tropical and Public Health Institute Sacarlal, Jahit ; Eduardo Mondlane University, Faculty of Medicine Madede, Tavares; Universidade Eduardo Mondlane Berté, Salimata; Centre Suisse de Recherches Scientifiques en Côte d'Ivoire Matsinhe, Graça; Ministério da Saúde Garba, Abdullahi; National Primary Healthcare Development Agency Brown, David; BCGI LLC / pivot-23.5°</p>
Primary Subject Heading:	Health services research
Secondary Subject Heading:	Global health, Health policy, Research methods
Keywords:	Human resource management < HEALTH SERVICES ADMINISTRATION & MANAGEMENT, Organisation of health services < HEALTH SERVICES ADMINISTRATION & MANAGEMENT, Protocols & guidelines < HEALTH SERVICES ADMINISTRATION & MANAGEMENT, Quality in health care <

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

	HEALTH SERVICES ADMINISTRATION & MANAGEMENT, STATISTICS & RESEARCH METHODS, PRIMARY CARE

SCHOLARONE™
Manuscripts



I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our [licence](#).

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which [Creative Commons](#) licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

Title

Does an innovative paper-based health information system (PHISICC) improve data quality and use in primary health care? Protocol of a multi-country, cluster randomised controlled trial in Sub-Saharan African rural settings.

Authors

- Xavier Bosch-Capblanch. Corresponding author. Swiss Tropical and Public Health Institute. Basel, Switzerland. And University of Basel, Basel, Switzerland. x.bosch@unibas.ch
- Angela Oyo-Ita. Department of Community Medicine, University of Calabar, Calabar, Nigeria. oyo_ita@yahoo.com
- Artur Manuel Muloliwa. Faculty of Health Sciences, Lúrio University, Nampula, Mozambique. muloliwa@yahoo.com.br
- Richard B Yapi. Centre Suisse de Recherches Scientifiques en Côte d'Ivoire, Abidjan, Côte d'Ivoire. And Centre d'Entomologie Médicale et Vétérinaire, Université Alassane Ouattara, Bouaké, Côte d'Ivoire. richard.yapi@csrs.ci
- Christian Auer. Swiss Tropical and Public Health Institute. Basel, Switzerland. And University of Basel, Basel, Switzerland. christian.auer@swisstph.ch
- Mamadou Samba. Ministère de la Santé et de l'Hygiène Publique, Abidjan, Côte d'Ivoire. And Université Félix Houphouët Boigny, Abidjan, Côte d'Ivoire. samba.mamadou@gmail.com
- Suzanne Gajewski. Swiss Tropical and Public Health Institute. Basel, Switzerland. And University of Basel, Basel, Switzerland. suzanne.gajewski@swisstph.ch
- Amanda Ross. Swiss Tropical and Public Health Institute. Basel, Switzerland. And University of Basel, Basel, Switzerland. amanda.ross@swisstph.ch
- L Kendall Krause. Bill & Melinda Gates Foundation, Seattle, USA. Kendall.Krause@gatesfoundation.org
- Nnette Ekpenyong. Department of Community Medicine, University of Calabar Teaching Hospital, Calabar, Nigeria. And Department of Community Medicine University of Calabar, Calabar, Nigeria. nnekon2015@gmail.com
- Ogonna Nwankwo. Swiss Tropical and Public Health Institute, Basel, Switzerland. And University of Basel, Basel, Switzerland. And Health Policy, Management and Economics Firm, Department of Community. Medicine, University of Calabar, Calabar, Nigeria. ogonna.nwankwo@swisstph.ch
- Anthonia Ngozi Njebuome. Swiss Tropical and Public Health Institute, Abuja, Nigeria. ngonjep@yahoo.com
- Sofia Mandjate Lee. Swiss Tropical and Public Health Institute; Manhiça, Mozambique. sofia_mandjate@yahoo.com.br
- Jahit Sacarlal. Faculty of Medicine, Eduardo Mondlane University, Maputo, Mozambique. jahityash2002@gmail.com
- Tavares Madede. Department of Community Health, Faculty of Medicine, Eduardo Mondlane University, Maputo, Mozambique. tmadede@gmail.com
- Salimata Berté. Centre Suisse de Recherches Scientifiques en Côte d'Ivoire, Abidjan, Côte d'Ivoire. sali.berte@csrs.ci
- Graça Matsinhe. Expanded Program on Immunization, Ministry of Health, Maputo, Mozambique. gmatsinhe@gmail.com.
- Abdullahi Bulama Garba. Director Planning, Research and Statistics, National Primary Health Care Development Agency, Abuja, Nigeria. ab.garba@nphcda.gov.ng.
- David W Brown. Managing Principal, BCGI LLC / pivot-23.5°, North Carolina, USA. david.brown@pivot235.org

Abstract

Introduction

Frontline health workers in remote health facilities are the first contact of the formal health sector and are confronted with life-saving decisions. Health information systems (HIS) support the collection and use of data. However, HIS focus on reporting and are unfit to support decisions. Since data tools are paper-based in most primary health care settings, we have produced an innovative paper-based HIS (PHISICC) using a Human Centred Design approach. We are carrying out a cluster-randomised controlled trial in three African countries to assess the effects of PHISICC compared with the current systems.

Methods and analysis

Study areas are in rural zones of Côte d'Ivoire, Mozambique and Nigeria. Seventy health facilities in each country have been randomly allocated to using PHISICC tools or to continuing to use the regular HIS tools. We have randomly selected households in the catchment areas of each health facility to collect outcomes' data. The baseline survey has been carried out in two of the three countries, the end-line survey is planned for mid-2021. Primary outcomes include data quality and use and coverage of health services and health workers satisfaction; secondary outcomes are additional data quality and use parameters, childhood mortality and additional health workers and clients experience with the system. Just prior to the implementation of the trial we had to relocate the studies in Mozambique and Côte d'Ivoire due to unforeseen logistical issues. The effects of the intervention will be estimated using regression models and accounting for clustering using random effects.

Ethics and dissemination

Ethics committees in Côte d'Ivoire, Mozambique and Nigeria approved the trials. We plan to disseminate our findings, data and research materials among researchers and policy makers. We aim at having our findings included in systematic reviews on health systems interventions and future guidance development on the matter.

Registration

Pan African Clinical Trials Registry - PACTR201904664660639. Registered 01/04/2019,
<https://pactr.samrc.ac.za/Search.aspx>.

Article summary

Strengths and limitations of the study

- These trials assess the effects of improving paper-based health information systems, which are greatly used particularly in remote, rural areas but which are neglected in research.
- The paper-based interventions have been developed using a Human Centred Design approach, with frontline health workers and designers driving the co-creation process.
- Despite the complexity of health systems interventions, we have applied robust experimental methods, together with qualitative research, to assess and understand the effects of the paper-based intervention. Robust evidence on health systems is more likely to gain the credibility of policy-makers and to make it into systematic reviews, guidance development and policy and practice.
- Research targeting frontline health workers in remote, rural areas has to take place where they live and work, which poses serious obstacles in the organisation, management and monitoring of the trials.
- These obstacles, aggravated by the COVID-19 pandemic, have challenged the mobility of the research team, the availability of the intervention in one of the countries and the duration of the trials.

Introduction

Frontline health workers (HW) in remote, rural health facilities (HF) in many countries are the first contact with the formal health sector of the population and they are confronted with life-saving clinical and public health decisions on a daily basis. Decisions are made by exerting a balanced judgment on the information related to health care events, such as making the correct diagnoses or deciding on which vaccinations a child should receive on a given day. In order to properly handle this information, appropriate data support tools and processes are required, referred to as the health information system (HIS); or Routine HIS or Health Management Information System [1]. In reality, though, HIS are primarily designed to report aggregated health events to the higher tiers of the health systems rather than to inform decision-making at the point of care [2].

Increasing pressure by donors and governments to collect more and more data has aggravated the situation, through the proliferation of data support tools that have overloaded frontline health workers compromising their capacity to deliver good quality of care and to delivery good quality data [3], for higher level decision-making.

Promising ‘quick fixes’, such as the scale up of digital HIS, are taking a long time to implement and face enormous challenges related to infrastructure, equipment and services necessary to run them. Besides, research evidence on the effects of digital solutions remains patchy and inconsistent, even in high-income country settings, where complaints about computerisation of clinical care have been raised [4,5]. Hence, it is very likely that paper tools will remain a primary, if not unique, data support mechanism particularly in remote, rural HF in many countries.

PHISICC (Paper-based Health Information System in Comprehensive Care) is a multi-year, multi-country, mixed-methods research project that aims at producing and testing an innovative paper-based HIS to improve data quality and use, decision making and health outcomes, at Primary Health Care (PHC). It is being carried out in selected areas within Côte d’Ivoire, Mozambique and Nigeria. The project started in 2015, producing a systematic review on the effects of HIS interventions [6,7]

1
2
3 and a framework synthesis on how HIS are understood in the literature in order to learn from past
4
5 experiences in HIS developments. This global evidence was coupled with studies to characterise the
6
7 existing HIS in the three countries, to understand how health workers interact with the HIS and to
8
9 identify entry points for HIS design improvements. With these bodies the research team was well
10
11 equipped to engage into a Human Centred Design (HCD) co-creative process with frontline HW to
12
13 design an innovative HIS (PHISICC). See Figure 1 for an illustration of the structure, processes and
14
15 evidence flow within PHISICC.
16
17

18
19 The impact of the PHISICC HIS on data quality and use, quality of health care and HW perceptions is
20
21 being assessed concurrently in rural areas in the three countries. We describe the design of the trial
22
23 here, consistent with CONSORT reporting guidelines [8] and the extension for cluster randomised
24
25 controlled trials (CRCT) [9]; see Additional file 1.
26
27

28 29 **Methods**

30 31 32 **Aim**

33
34 The aim of the trial is to address the research question: what are the effects of an innovative paper-
35
36 based HIS (PHISICC) on data use and quality, quality of health and HW perceptions compared with
37
38 the current HIS, in rural PHC settings?
39
40
41

42 43 **Patient and public involvement**

44
45 There was no public or patient involvement in the design of the study or selection of study areas
46
47 because the intervention being assessed in these trials target health care providers and decision-
48
49 makers, rather than patients or the public in general. Population in the catchment area of selected
50
51 health facilities, potentially using their services, were only approached in order to assess the studies
52
53 outcomes.
54
55

56
57 On the other hand, we have involved health systems stakeholders and frontline health workers.

58
59 Ministries of Health at several levels participated in the preparation of the research proposal
60

1
2
3 (personal consultations), in the characterisation of health information systems that preceded the
4 trials (countries workshops), and throughout all project components (additional workshops,
5 newsletters and personal communication). Frontline health workers in the three countries have co-
6 created the intervention (i.e. paper based tools) through workshops, personal feedback and piloting
7 under real live conditions. Some of them are part of the research team and co-authoring this
8 manuscript.
9
10
11
12
13
14
15
16

17 **Study design**

18
19 The study is a CRCT in each of the three countries. In each setting, 70 health facilities are randomised
20 to intervention or control (35 per arm). The intervention arm HF use the new PHISICC tools
21 (substituting the usual HIS tools) and the control arm HF use the regular HIS tools. The trial is
22 implemented in the real life contexts of HF carrying out their usual duties.
23
24
25
26
27
28

29 The trials started between the end of 2019 and beginning of 2020, depending on the country, when
30 the intervention was installed and the baseline surveys carried out. Data collection will last until mid-
31 2021.
32
33
34
35
36

37 **Study areas**

38
39 Ministries of Health (MOH) officials in several countries were contacted before submitting the
40 proposal to the funding agency in order to explore the willingness to engage in a project focusing on
41 paper-based tools. Officials in several countries rejected the offer on the grounds of upcoming
42 digitalisation plans of the HIS in the country. We partnered with MOH that found the research
43 relevant to their context in three countries.
44
45
46
47
48
49

50 In each country, the eligibility criteria of study areas were that they had to belong to the operational
51 area of research partners; contain a large enough number of health facilities and their catchment
52 population; include vulnerable population (e.g. with low vaccination coverage, high childhood
53 mortality); and be comparatively neglected in terms of infrastructure and services. We excluded
54 areas with concurrent research or other types of activities that could conflict with the CRCT (such as
55
56
57
58
59
60

1
2
3 the co-existence of another health-related study, massive developments in infrastructure or activities
4 involving migration of the population, such as temporary work sites or changes in working sites) and
5 areas with threats to safety or security that could jeopardise research activities.
6
7
8

9
10 The study areas are located in Adzopé, Agboville, Tiassalé and Sikensi districts (Côte d'Ivoire); in
11 Funhalouro, Govuro, Homoine, Inhambane, Inharrime, Inhassoro, Mabote, Maxixe and Panda
12 (Inhambane province, Mozambique); and in Yala Local Government Authority (Cross River State,
13
14
15
16
17
18 Nigeria).

19 **Eligibility of health facilities**

20
21 The intervention is implemented at the HF level. The eligibility criteria of the HF were that they had
22 to be located in the study areas, belong to the governmental health sector and their main activity
23 should be the delivery of PHC services. HF were excluded if they had specialised clinical services,
24 inpatients, physicians providing care or with plans for staff turn-over involving intervention and
25 control HF.
26
27
28
29
30
31

32
33 A 'master list' of eligible health facilities was prepared based on information provided by the MOH
34 across all study areas. We aimed at selecting 70 of the eligible HF in each country, using simple
35 random techniques in R [10]. See in Figure 2 the selection and allocation trial flow chart.
36
37
38
39

40 **Allocation and blinding**

41
42 Allocation of the 70 HF per country into the intervention and control arms took place in a formal
43 event, gathering research partners and MOH officials to offer transparency and promote study
44 ownership by local and national authorities. Equally sized, folded pieces of paper with the names and
45 codes of included HF written on them were introduced in an opaque receptacle where they were
46 manually and blindly mixed. A second receptacle contained two equally sized pieces of paper, one
47 with the word 'intervention' and another one with the word 'control'. A selected person in the
48 meeting, not belonging to the research team, extracted one piece of paper at a time to reach half the
49 number of included HF. Then, a paper was extracted from the second receptacle to assign those HF
50
51
52
53
54
55
56
57
58
59
60

1
2
3 to the intervention or control arms. The rest of the papers were extracted as well to verify
4
5 completeness and no duplication of names, and those HF assigned to the other arm.
6

7
8 Once HF were selected, all villages or settlements for each health facility catchment area were listed
9
10 and three in each catchment area were selected. In practice, we selected all villages because the
11
12 numbers were below (in Côte d'Ivoire) or just above (in Nigeria) the needs. For each village, we used
13
14 Google satellite maps to identify and geo-locate every visible roof. Where there were many houses
15
16 per village (roughly, more than fifty or so), a researcher would mark four points in the map slightly
17
18 beyond the northernmost, southernmost, easternmost and westernmost roofs seen and 30 random
19
20 points were selected within that square. From the mapped points, 10 per village (with 10 more acting
21
22 as reserve) were randomly selected and marked on another map used in the field for data collectors
23
24 to approach households. Where technical problems impeded this approach in a given village, a field
25
26 supervisor would rotate a bottle on the floor towards the centre of the village and would select at
27
28 random 10 households in the direction pointed by the bottle, from the outer limit of the village till
29
30 the centre [11].
31
32
33

34
35 Blinding is only feasible for the research team members carrying out the CRCT data collection and the
36
37 analyses of the CRCT findings. The intervention (i.e. paper tools) are by design very different from the
38
39 existing system and it is not possible to blind participants or principal researchers.
40

41
42 We already had the agreement of the MOH and selected HF compliant with the inclusion criteria
43
44 were provided with the intervention shortly after completing the baseline data collection. Therefore,
45
46 recruitment as such took place at the same time of the allocation of HF into intervention and control
47
48 arms.
49

50 51 **The intervention**

52
53 The PHISICC paper-based intervention is a full set of paper-based tools to support decision-making by
54
55 frontline HW. These are the only tools to be used by HW in the intervention arm. The PHISICC tools
56
57 encompass the whole system (i.e. recording and reporting) and all clinical and public health care
58
59
60

1
2
3 areas and are characterised by: a common visual language (e.g. spaces for digits and text),
4
5 standardised formats across health care areas; support to critical data items (e.g. respiratory rate in
6
7 infants); graphic artefacts to distinguish severity degrees of signs or symptoms; documentation of
8
9 diagnoses and treatment decisions; and aides memoires, among others.

10
11
12 The PHISICC tools have been developed from May 2017 till June 2019, including production, using a
13
14 Human Centred Design approach [12]. A strength of the Human Centred Design approach is its ability
15
16 to unlock the user's perspective so that designers can build solutions that are fully reality-based and
17
18 work well. Co-creation groups were formed in each country with researchers, staff from partner
19
20 institutions and healthcare workers, led by a team of professional designers. Research team
21
22 members supervised and coordinated exclusively the feedback on the contents of the tools, to
23
24 ensure compliance with each country clinical guidelines. At the outset of the process, the design
25
26 focused on three health care areas (i.e. antenatal care, vaccination and sick child) and slowly
27
28 extended the new visual language to other health care areas. Initial workshops served to brainstorm
29
30 on problems and potential design solutions, without any other rule than being comprehensive and
31
32 not rejecting a single idea. Designers, then, formalised some of the most promising solutions and a
33
34 first round of exchanges within the co-creation team was used to address misinterpretations or
35
36 inconsistencies. There were two in-the-field testing rounds in Mozambique, two in Côte d'Ivoire and
37
38 three in Nigeria and uncountable exchanges through teleconferences and email, in-between. The
39
40 prototypes were considered final when no errors were detected, were compliant with data needs in
41
42 each country and comments from the field could not be accommodated in the design concept or
43
44 there was no consensus on minor or formal issues being raised.

45
46
47 The PHISICC tools have been produced in French for Côte d'Ivoire, in Portuguese for Mozambique
48
49 and in English for Nigeria, which are the official languages used in the health systems in the three
50
51 countries. They include the official logo of the MOHs. Health care areas covered include: family
52
53 planning, antenatal care, including tetanus toxoid vaccination, delivery, post-natal care, vaccination,
54
55
56
57
58
59
60

1
2
3 sick child, adults outpatient consultation, tuberculosis diagnosis and treatment, and HIV. Referral
4
5 forms were also designed.
6

7
8 The PHISICC tools have three sub-components: registers, tallies and reports. Registers are formed by
9
10 seven DIN-A3 and one DIN-A4 (for referrals) book covering all health care areas except for
11
12 tuberculosis treatment, for which DIN-A3 cards were used. Register books have 100, 200 or 400
13
14 pages depending on the country and health care area. They are used to record individual clients' data
15
16 for each health care event, either of clinical or public health nature. Some register books have clinical
17
18 notes at the very beginning, as 'aide memoires', and an example of a filled-in form, to assist HW
19
20 when doubting how to proceed.
21
22

23
24 Tallies are DIN-A3 single sheets which contain a list of the indicators to be transferred to higher levels
25
26 of the health system, with a series of small ovals, grouped in fives, to mark with tally sticks with a
27
28 pen. In contrast to the current systems that have no tallies or only for vaccination, tallies were
29
30 created for all health care areas. In the middle-right side of the tally, a column accommodates cells
31
32 aligned with the ovals to insert the count for each indicator; and in the far right of the sheet there is
33
34 a replica of the count column, separated with a perforated line, which is detached and sent, as part
35
36 of the monthly report to the higher level in the health system.
37
38

39
40 While current HIS tools are consistently organised in tabular formats and books, where each clinical
41
42 event is recorded in a row and each variable (e.g. age, gender, HIV status, diagnosis) in a column,
43
44 PHISICC tools incorporated several innovations; in summary: a visual language to guide the clinical
45
46 decisions of health workers based on severity (i.e. if it is recorded that a child has convulsions, a
47
48 visual artefact indicates severity), more space for clinical data (e.g. vital signs), inclusion of all critical
49
50 information to assess patients (e.g. obstetric history, gestational age, fundus height, breath rate in
51
52 infants), consolidation of information of all antenatal care visits for a single pregnant woman in the
53
54 same page, among many other formal and contents improvements, including improved aesthetics.
55
56

57
58 We aimed at creating "a system" (not just some tools) focusing on decision making by frontline
59
60 health workers. The epidemiological and public health contexts in the three countries are similar, as

1
2
3 confirmed by the similarities in the existing HIS between the three. The visual language and the
4
5 recording forms were common to the three countries because clinical decisions are common to the
6
7 three; although forms allowed for specific tests or treatments. The reporting component was
8
9 adapted to each country set of indicators, although the visual language and reporting processes were
10
11 harmonised.
12

13
14 During three or four days, HW were trained on HIS before the start of the trial. In the intervention
15
16 arm they were trained on the PHISICC tools; and the control arm received a refresher training about
17
18 the regular tools, during the same number of days.
19

20
21 Additionally, given that the regular tools already contained information on past vaccination history of
22
23 children still to complete their vaccination schedule, we created a mechanism to retrieve data of
24
25 children's vaccination status to transcribe into the new vaccination register book in the intervention
26
27 arm ('system transition').
28

29
30 Tools were endorsed by MOH, printed in local printing companies and distributed to HW at the end
31
32 of the training sessions. A digital spreadsheet was created to monitor consumption and order
33
34 additional tools to cover health facility needs during the life of the trial.
35
36

37 38 **Outcomes**

39
40 There are five primary outcomes (Table 1). Vaccination adherence is defined as the total number of
41
42 vaccine doses given in the correct time interval to children in households in the health facilities
43
44 catchment villages of those over the total number of vaccine doses that should have been given
45
46 during the same period. Antenatal care visits uptake will also be considered depending on the
47
48 expected number of pregnancies in the study areas. Both are used as proxies for health outcomes in
49
50 terms of protection against disease [13] and prevention of pregnancy complications [14] and are
51
52 assessed in a random sample of households in the health facilities catchment areas. Data
53
54 concordance is defined as the level of agreement of HIS indicators between (i) records of health care
55
56 events (re-counts), (ii) tallies (re-counts) and (iii) reports (aggregated data to higher levels of the
57
58
59
60

1
2
3 system) [3]. The appropriateness of treatment will be measured using the diagnostics scope in the
4
5 sick child (i.e. number of different diagnoses per sick child consultation); and treatment
6
7 appropriateness (i.e. number of prescribed treatments that are supported by a documented
8
9 diagnosis). These outcomes will be assessed in a random sample of records and corresponding
10
11 reports during the last four months of the study period. Health workers satisfaction will be assessed
12
13 in all health workers in included health facilities using a standardised questionnaire [15,16,17].
14

15
16 While the intervention targets HF, some of the outcomes are measured at the level of HF, and some
17
18 from patients clustered within HF catchment areas.
19

20
21 Secondary outcomes are classified under the following domains: data quality, data use, mortality,
22
23 HW experience, client experience and resource consumption:
24

- 25
26 • Data quality, assessed in a sample of records
 - 27
28 ○ Completeness of recording and reporting in specific forms; i.e. prevalence of unduly
29
30 missing data items; partograph used;
 - 31
32 ○ accuracy of recorded figures in comparison to real events (e.g. physical counting of
33
34 commodities, such as number of 500mg Paracetamol tablets as recorded versus
35
36 number of 500mg Paracetamol tablets as counted);
 - 37
38 ○ timeliness of reporting, as documented by time stamps in forms;
 - 39
40 ○ loss of data or data which does not reach the next upper administrative level.
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

- Data use, assessed in a sample of records
 - in terms of knowledge (e.g. vaccines due based on date of birth; weight for length assessments);
 - cases of different conditions properly treated in (e.g. diarrhoea cases given oral rehydration therapy according to national guidelines; pneumonia cases given appropriate antibiotic according to national guidelines;
 - public health decisions: availability of lost to follow up lists or plans for vaccination, tuberculosis and or HIV/AIDS treatment control;
 - occurrence of stock outs of essential drugs.
- Overall under-5s mortality and under-5s mortality excluding peri-natal mortality [18], in a sample of households in health facilities catchment areas.
- Health workers' 'human experience' and satisfaction (all health workers).
- District health information officers' 'human experience' (selected health care programme managers).
- Clients' 'human experience' and satisfaction, in a sample of households in health facilities catchment areas.
- Resources consumption (e.g. time use, costs)
 - intervention costs: tools, training, start-up;
 - time used for recording and reporting (e.g. time-motion study) [19];
 - cost-effectiveness per unit of additional improvement in outcomes of interest.

It is worthwhile to note that outcomes that do not relate to data quality and use will be assessed using additional data collection tools (e.g. survey questionnaires), which are the same for intervention and control health facilities. Hence, the effects of the intervention cannot be attributed to the changes in performance of the paper tools routinely used to record health care events in intervention and control health facilities, which are different by design.

1
2
3 In addition, we will consider ‘explanatory outcomes’ that will help to understand how the measured
4 effects have taken place and why. We will look at the details of the interplay between the
5 intervention, the system, the users and the context. Process indicators will be based on the
6 documented activities that have taken place, from the conception of the intervention, up to its
7 implementation, monitoring and evaluation. Process indicators may include: intervention set up and
8 implementation, monitoring of the use of the intervention, special activities targeted at vulnerable
9 populations, district reactions related to the intervention, handling of data coming from the new
10 system, sustainability based on costs information and perceptions, alignment with national health
11 policies and donor priorities. We will also explore health care services characteristics looking at
12 generic indicators from health facilities, such human resources profiles and relations with the
13 communities, population characteristics and system and context characteristics captured in early
14 stages of the project, where data are available.

31 **Sample size calculations**

32
33 The required sample sizes for each primary outcome were determined using simulation to
34 incorporate the clustering easily (Table 1) and to take the baseline and end-line surveys into account.
35 Briefly, we simulated 1000 trials with variation between them caused by drawing different samples
36 from the same distributions. We then used the regression models detailed in the data section to
37 analyse each of the simulated trials and estimate the power as the proportion of trials which
38 detected the effect of the intervention as significant. The simulation code was written in R
39 (supplementary files 1 and 2).

40
41 For each country, we require the probability of α , a type I error (rejecting the null hypothesis when it
42 is actually true) to be less 0.05 and the power to be at least 80%.

43
44 For vaccination adherence, using a sample size of 35 HF per arm, we would have 80% power in each
45 country to detect as significant a difference between a proportion of due vaccines given from 75% in
46 the control to 85% in the intervention arm, assuming one child per household, 30 households per HF
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 and a between-HF variation equivalent to a k of 0.25, where k is equal to the standard deviation
4
5 divided by the mean. The value of k is unknown, but was chosen in line with general observations by
6
7 Hayes and Bennet [20].
8

9
10 For data quality outcomes, with 35 HF per arm we would be able to detect as significant a difference
11
12 from a ratio of 0.7 (reported : recorded) vaccinations in the control arm to 0.8 with the intervention
13
14 with 80% power, assuming 100 recorded vaccinations per HF and a standard deviation of 0.25 in the
15
16 ratios between HF.
17

18
19 In terms of diagnostic scope, we would be able to detect an increase in the proportion of child-visits
20
21 with more than one diagnosis from 30% to 35% with 80% power with 35 HF per arm, 60 records per
22
23 HF and assuming a k of 0.25.
24

25
26 We would be able to detect as significant an increase from 50% of treatments having a
27
28 corresponding appropriate diagnosis to 60% with 80% power assuming 35 HF per arm, 1 treatment
29
30 per child, 25 children per HF and variation between HF corresponding to $k = 0.25$.
31

32
33 For the outcome related to health workers' satisfaction, we would be able to detect as significant an
34
35 increase from 50% of health workers satisfied to 75%, with 80% power assuming 35 HF, three health
36
37 workers per HF and a variation between HF equivalent to $k = 0.25$. Since this variable is measured in
38
39 the end-line survey only, we used the formula in Hayes and Bennet [20].
40

41
42 In summary, in each country we require 35 HF per arm, three HW per HF, 100 vaccination records
43
44 per HF, 60 sick child records per health facility and 30 children per health facility catchment area.
45
46
47

48 **Data collection and management**

49
50 Data collection took place at baseline and will take place again at the end of the study. Data is
51
52 collected from health facilities, from the households in the catchment areas of the included health
53
54 facilities and also from district offices.
55

56
57 For data quality and data use outcomes, HF registers, tallies and reports will be scrutinised. For
58
59 population based outcomes, we carry out household surveys at baseline and at end-line. We use
60

1
2
3 standard approaches for these types of surveys [21]. Households are visited, the research project is
4
5 briefly introduced and consent requested. Ideally, mothers of alive children or women in child-
6
7 bearing age were interviewed in order to obtain information on living children (i.e. vaccination
8
9 history) and death events, respectively, using home-based records if available and accessible.
10
11 Patients' satisfaction will be assessed using the PSQ-18 satisfaction questionnaire [22,23,24].
12
13
14 Essentially, the tool enables practitioners to investigate the extent to which their health care service
15
16 meets the perceived needs of their client group and pinpoint areas for improvement [24]. The
17
18 interview will be conducted with consenting patients as close to their care encounter as possible
19
20 [25]. Data tools are translated into the official languages of the study countries and pilot tested for
21
22 consistent meaning and relevance to the setting. Data collectors are also able to communicate in
23
24 local languages. The Satisfaction of Employees in Health Care (SEHC) survey is a validated tool to
25
26 assess staff satisfaction. It was first developed and validated in a low-income country (Ethiopia) [26]
27
28 and later successfully validated in a high-income country (USA) [27].
29
30

31
32 We use a mix of paper and electronic data (ODK [28]) collection tools. Data collectors are trained to
33
34 minimise error. Tools are piloted before implementing. ODK data is regularly stored and sent to
35
36 secure servers, as soon as data collectors reach their office base. Data from paper tools is double
37
38 entered and compared and sent to secure servers. Each data collection tool has its corresponding
39
40 electronic database that is cleaned and submitted to the analyses. All data is anonymised at the point
41
42 of data collection or as soon as possible in the data management process. Data is labelled with an
43
44 arm code (e.g. 'A' or 'B') without any further information allowing to disclose which data items
45
46 belong to the intervention or to the control arms, ensuring blinding during data analyses.
47
48

49
50 Quality will be assured through several mechanisms: piloting of data collection tools; thorough
51
52 training of field workers; checking missing data related; double, independent data entry from papers
53
54 into digital databases; early descriptive analyses to detect potential outliers; fieldworkers tracking
55
56 and supervision.
57
58
59
60

Data analysis

The analysis will be carried out for each country separately, and based on intention-to-treat.

Baseline population and health facility characteristics (i.e. basic demographic characteristics of population and health workers, professional profile of health workers, health facility size and services) will be summarised. If large imbalances are observed at baseline, the variables can be used to adjust the effect estimate comparisons [29,30].

The analyses vary for the different primary outcomes due to the unit of measurement and levels of clustering, the type of variable, and whether measurements were taken at baseline and endpoint or endpoint only. We use regression models to allow us to estimate the effect of the outcome while flexibly accounting for these issues and allowing adjustment for potential confounders.

Logistic regression will be used for the binary variables: vaccine adherence is measured by determining whether each vaccine due was received, and treatment appropriateness by whether each treatment was correctly prescribed. Data concordance and diagnostic scope are count variables and may be analysed with Poisson regression, depending on their distribution. The regression model for HW satisfaction will depend on how it is distributed.

The outcomes have different levels of clustering (children or consultations, HW, HF). We will account for these levels of clustering by including random effects in the regression models.

Four of the primary outcomes are measured at baseline and end-line. The effect of the intervention will be estimated using an interaction term between arm and survey in the regression models: i.e. is the change in the outcome between baseline and follow-up in the intervention arm different to the change between baseline and follow-up in the control arm? The effect of HW satisfaction, measured only at end-line, will be estimated as the difference between the intervention and control arm.

All estimates for the effect of the intervention will be presented with 95% confidence intervals. The analyses will be carried out using R [31].

Measures to minimise bias

Statistical analyses will be carried out blindly, without knowledge of what health facilities or population in the catchment area belong to the intervention or control groups. Only when the analysis code is considered as definitive and fixed, will results be shared with the wider investigators team and the arms for health facilities and population will be disclosed.

Outcome measurement bias may take place where data from the HIS, which is the focus of the intervention, is used to measure outcomes. However, we will minimise this by assessing population based outcomes at household level.

Contamination (i.e. the intervention affects individuals or units assigned to the control arm) may happen via the exchanges between health workers from health facilities in both arms; for example: in monthly district data quality meetings, managerial meetings; or through inputs from supervisors who influence control health facilities with intervention tips encountered in health facilities of the intervention arms. One mechanism to address this issue is using a district-based cluster randomisation scheme. However, we consider that (i) contamination, despite increasing the awareness of health works in control health facilities, will hardly influence the decision making mechanisms that the HIS intervention focuses on; and (ii) randomisation at the level of district poses additional challenges that are not worth the marginal benefit of reducing a doubtful contamination [32].

The spill-over effect (i.e. benefits of the intervention extend beyond their direct recipients) [33] may take place in higher levels of the health systems; e.g. district data managers and programme managers may experience the benefits of better structured and more timely data produced in health facilities in the intervention arms. The trial will have no capacity to quantitatively account for spill overs at higher levels of the system, due to the limited number of higher level administrative areas that will be involved in the trial. However, through process indicators, we will consider potential benefits and harms of the intervention at higher levels of the system.

1
2
3 A challenge is the Hawthorne effect (i.e. observer effect). Both health workers in the intervention
4 and in the control sites will have an awareness of being observed as data collection activities will be
5 at the same level of intensity in the two arms. Therefore, there should be no differential effect.
6
7

8
9
10 Analyses will be based on the intention-to-treat. It is important to closely monitor if the intervention
11 HFs consistently use the new HIS tools and approaches. The data collection team and the trial
12 monitoring team will check if old forms are still being used in the intervention health facilities.
13
14

15
16 However, we do not expect health facilities to migrate between intervention and control arms, or
17 vice versa, due to feasibility issues. On the other hand, some household members in a given
18 catchment area may decide to seek for health care in a health facility belonging to another trial arm.
19
20
21

22 In these cases, households will be analysed as belonging to the original trial arm.
23
24
25

26 **Ethics and dissemination**

27
28
29 Ethics committees in Côte d'Ivoire, Mozambique and Nigeria approved the study in their respective
30 countries. To date, some modifications to the protocol have taken place. In Côte d'Ivoire, we decided
31 to select study areas close to the research institution base on logistics and practical reasons, instead
32 of selecting an area in the north of the country, where poorer health indicators have been described.
33
34

35
36 In Mozambique, the low density of HF per population implied extremely vast distances between HF
37 and this, coupled with the rainy season, made the trial unfeasible in the originally selected Nampula
38 province. After consultations, we decided to move the trial to the province of Inhambane and cancel
39 the household survey. The allocation of HF to the intervention and control arms was completed using
40 random number generation.
41
42
43

44
45 We plan to disseminate the findings of the trials as one of the few examples of studies assessing the
46 effects of health information systems interventions using experimental study designs [34]. Most of
47 the experimental studies on HIS are circumscribed to specific health care areas (e.g. tuberculosis,
48 vaccination, cardiovascular disease) and very few have a system-wide approach (e.g. PHC) [34].
49
50
51

52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

1
2
3 limited capacity to provide reliable explanations of complex health system issues [35]. While we
4
5 acknowledge these limitations, there is also a need for more robust evidence on the effects of these
6
7 types of health system interventions [36] and there are also good examples of experimental studies
8
9 reporting findings that can make it to the policy arena [37]. When embarking on this research, we
10
11 considered from the outset the type of evidence required to be disseminated and included into
12
13 systematic reviews [38], guidance development [39] and eventually recommendations for policy and
14
15 practice [40].
16
17
18
19
20

21 We acknowledge the challenges of carrying out research on health care provided to remote, rural
22
23 communities (in this case in Sub-Saharan Africa). However, it is only in these remote areas where
24
25 research about their specific problems and needs can take place. Challenges included long distances,
26
27 poor conditions of roads, unreliable communications and limited food and accommodation services,
28
29 all of them to be proactively handled to keep the quality of work and the morale of researchers and
30
31 collaborators. We expect that the dissemination of findings in meetings, conferences and
32
33 publications will contribute to a better understanding of what it takes to make research in
34
35 challenging contexts.
36
37
38

39 The engagement and ownership of partners within this research has also been instrumental in order
40
41 to plan and implement the CRCT. The intervention actually targets a governmental sub-system (the
42
43 HIS) for which we required more than permission but also endorsement and active support. We
44
45 achieved this level of collaboration by ensuring the participation of key stakeholders in key phases of
46
47 the whole project, from inception till the implementation of the last phases, through frequent
48
49 communication and workshops. The PHISICC programme includes targeted activities to keep
50
51 decision-makers engaged and we are planning to share the findings through workshops as well as
52
53 online and face-to-face events to disseminate the lessons learned from the trial and the whole
54
55 research.
56
57
58
59
60

1
2
3 We also expect that the dissemination of our findings among partners and competitors will
4
5 contribute to the current debates on the digitalisation of health information systems. WHO
6
7 recommendations on the matter are weak since the underlying evidence to support these
8
9 recommendations is inconsistent [41]. The principles and methodological approaches in PHISICC can
10
11 be applied to the development of any technological solution, being on paper, digital or mixed.
12
13
14 Finally, we expect that the results of the trials, both quantitative and qualitative, will be able to
15
16 inform policies on how to make HIS responsive to providers' decision-making needs, particularly in
17
18 health services where the most vulnerable live.
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Authors' contributions

XBC, AOI, AM, RBY and CAU prepared the proposal for the funding agency, conceived the study and produced the data collection tools. SG ensured the regulatory, ethical and trial monitoring components. AR developed the analytical approaches and made the sample size calculations. RBY, MS and SB adapted the protocol to the context of Côte d'Ivoire and managed the administrative and ethical approvals in the country; AOI, NE, ON, ANN and ABG likewise in Nigeria; AM, SML and GM, in Nampula province (Mozambique); JS and TM adapted the protocol and acquired ethical and administrative clearances for Inhambane province (Mozambique). DB is chair of the PHISICC Technical Advisory Group (TAG) and has coordinated multiple formal and informal inputs. LKK and DB have advised on the adequacy of the study protocol within the overall PHISICC proposal and TAG advices. All country teams participated in PHISICC workshops and ensured that the protocol was suitable to countries realities; developed data collection tools and training materials. They are responsible for the implementation of the trial in each country. XBC drafted the first version of the manuscript. All authors commented on several versions of the manuscript.

Acknowledgements

Members of the PHISICC Technical Advisory Group, chaired by David Brown: Blanche Anya (WHO AFRO), Abdallah Bchir (former Gavi), Marta Gacic-Dobo (WHO HQ), Richard Greffé (AIGA), Pamela Mitula (WHO AFRO), Sandy Oliver (UCL Institute) and Chris Wolf (BMGF).

Research collaborators: Momade Ali, Celso Belo, Bassirou Bonfoh, Lisa Diallo, John Ferreira, Bernard Guessanbi, Caitlin Jarrett, Inza Koné, Felix Malé, Kouadio M'bra, David O'Donnell, Damaris Rodríguez (Sonder Collective), Melanie Wendland and Meike Zuske (Swiss TPH).

Stakeholders from Côte d'Ivoire, Mozambique and Nigeria, participating in the consultation processes.

Data sharing statement

We report on a protocol of three Cluster Randomised Controlled Trials and therefore no data is available yet. On completion of the trials, access to data will be available from the national research institutions in Côte d'Ivoire, Mozambique and Nigeria, in publications and in the funding agency.

Individual participants' data on vaccination and antenatal care outcomes as well as health workers and users perceptions will be anonymised and made available via an online data repository for any purpose and access will be granted following a review of requests by the Swiss TPH contract officer. Data, with DOIs, will be made available during the second semester of 2021.

Available documents include study protocols, analytical plan, informed consent forms and analytical code.

Ethical approvals in Côte d'Ivoire, Mozambique and Nigeria

- Comité National Ethique des Sciences de la Vie et de la Santé (CNESVS), reference: 024-19/MSHP/CNESVS-kp (Côte d'Ivoire)
- Comité Institucional de Bioética para Saúde da Universidade Lúrio, reference: 16.2/Julho/CBISUL/19 (Mozambique)
- Secretary, Government of Cross River State of Nigeria, Ministry of Health, Calabar Health Research Ethics Committee, reference: CRS/MH/HREC/018/Vol. V1/151 (Nigeria)
- Ethikkommission Nordwest- und Zentralschweiz (EKNZ), reference: 2018-01059 (Switzerland).

Funding

Funded by the Bill & Melinda Gates Foundation, grant number INV-010193 / OPP1135947.

Abbreviations

BMGF Bill & Melinda Gates Foundation

CRCT Cluster Randomised Controlled Trial

sd Standard deviation

HF Health Facility

HW Health worker

WHO World Health Organisation

Additional files

- Additional file 1: CONSORT statement checklist.

Competing interests

No, there are no competing interests for any author.

References

- 1 WHO. Monitoring results with health facility information systems: A technical consultation Glion-sur-Montreux, Switzerland, 11–12 June 2014. WHO 2014.
- 2 WHO. Health facility & community data toolkit. WHO 2014.
- 3 Bosch-Capblanch X, Ronveaux O, Doyle V, Remedios V, Bchir A. Accuracy and quality of immunization information systems in forty-one low income countries. *Trop Med Int Health*. 2009 Jan;14(1):2-10. doi: 10.1111/j.1365-3156.2008.02181.x. PMID: 19152556.
- 4 Loder E. Promises, promises. *BMJ*. 2017 Mar 23;356:j1446. doi: 10.1136/bmj.j1446. PMID: 28336513.
- 5 Wenger N, Méan M, Castioni J, Marques-Vidal P, Waeber G, Garnier A. Allocation of Internal Medicine Resident Time in a Swiss Hospital: A Time and Motion Study of Day and Evening Shifts. *Ann Intern Med*. 2017 Apr 18;166(8):579-586. doi: 10.7326/M16-2238. Epub 2017 Jan 31. PMID: 28135724.
- 6 Bosch-Capblanch X, Zuske M, Auer C, Amadasoun S. Effects of interventions to improve ‘health information use systems’ on the quality and use of information. National Institute for Health Research. PROSPERO - https://www.crd.york.ac.uk/prospero/display_record.php?RecordID=35218.
- 7 Bosch-Capblanch X, Zuske M, Auer C. Making information support systems relevant to decision-makers: an effectiveness systematic review and a framework synthesis. Abstracts of the 25th Cochrane Colloquium, Edinburgh, UK. *Cochrane Database of Systematic Reviews* 2018;(9 Suppl 1). <https://doi.org/10.1002/14651858.CD201801>.
- 8 Schulz KF, Altman DG, Moher D, for the CONSORT Group. CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials.
- 9 Campbell MK, Piaggio G, Elbourne DR, Altman DG; for the CONSORT Group. Consort 2010 statement: extension to cluster randomised trials. *BMJ*. 2012;345:e5661.
- 10 R Core Team. R: A Language and Environment for Statistical Computing [Internet]. Vienna, Austria; 2020. Available from: <https://www.R-project.org/>
- 11 Bostoen K, Chalabi Z. Optimization of household survey sampling without sample frames. *Int J Epidemiol*. 2006 Jun;35(3):751-5. doi: 10.1093/ije/dyl019. Epub 2006 Feb 15. PMID: 16481364.
- 12 Johnson, Sandhu, Tyler, “The Next Step for Human-Centered Design in Global Public Health” *Stanford Social Innovation Review*, Oct. 22, 2019.
- 13 WHO. Global Vaccine Action Plan 2011-2020. WHO 2013.

- 1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
- 14 Carroli G, Villar J, Piaggio G, Khan-Neelofur D, Gülmezoglu M, Mugford M, Lumbiganon P, Farnot U, Bergsjø P; WHO Antenatal Care Trial Research Group. WHO systematic review of randomised controlled trials of routine antenatal care. *Lancet*. 2001 May 19;357(9268):1565-70. doi: 10.1016/S0140-6736(00)04723-1. PMID: 11377643.
- 15 Alpern R, Canavan ME, Thompson JT, McNatt Z, Tatek D, Lindfield T, Bradley EH. Development of a brief instrument for assessing healthcare employee satisfaction in a low-income setting. *PLoS One*. 2013 Nov 5;8(11):e79053. doi: 10.1371/journal.pone.0079053. PMID: 24223878; PMCID: PMC3818514.
- 16 Doll WJ and Torkzadeh G. The measurement of end-user computing satisfaction. *MIS Quarterly* 1988;12(2):259–74; 19.
- 17 Ilias A, Yasoa MR, Razaka MZA and Rahmana RA. The study of end-user computing satisfaction (EUCS) on computerised accounting system (CAS) among Labuan F.T. Government sectors: a case study in the responsibility centres. *Labuan e-Journal of Muamalat and Society* 2007;1:1–13.
- 18 Arifeen SE, Hoque DM, Akter T, Rahman M, Hoque ME, Begum K, Chowdhury EK, Khan R, Blum LS, Ahmed S, Hossain MA, Siddik A, Begum N, Sadeq-ur Rahman Q, Haque TM, Billah SM, Islam M, Rumi RA, Law E, Al-Helal ZA, Baqui AH, Schellenberg J, Adam T, Moulton LH, Habicht JP, Scherpbier RW, Victora CG, Bryce J, Black RE. Effect of the Integrated Management of Childhood Illness strategy on childhood mortality and nutrition in a rural area in Bangladesh: a cluster randomised trial. *Lancet*. 2009 Aug 1;374(9687):393-403. doi: 10.1016/S0140-6736(09)60828-X. PubMed PMID: 19647607.
- 19 Tilahun H, Fekadu B, Abdisa H, Canavan M, Linnander E, Bradley EH, Berman P. Ethiopia's health extension workers use of work time on duty: time and motion study. *Health Policy Plan*. 2017 Apr 1;32(3):320-328. doi: 10.1093/heapol/czw129. PubMed PMID: 27658649.
- 20 Hayes RJ, Bennett S. Simple sample size calculation for cluster-randomized trials. *Int J Epidemiol*. 1999 Apr;28(2):319-26.
- 21 The DHS program. Demographic and Health surveys. <https://dhsprogram.com/>.
- 22 Thayaparan, A. a. (2013). The Patient Satisfaction Questionnaire Short Form (PSQ-18) as an adaptable, reliable, and validated tool for use in various settings. *Med Educ Online*, 18.
- 23 The Health Foundation. (2013, June). Measuring patient experience (no.18). Retrieved Nov 6, 2018, from www.health.org.uk: <https://www.health.org.uk/publication/measuring-patient-experience>.

- 1
2
3
4
5
6
7 24 Grogan, S. C. (2000). Validation of a questionnaire measuring patient satisfaction with general
8 practitioner services. *Quality in Health Care*, 9, 210-215.
- 9
10 25 LaVela, S. L. (2014). Evaluation and measurement of patient experience. *Patient Experience*
11 *Journal*, 1(1), Article 5.
- 12
13 26 Alpern, R. C. (2013, November). Development of a Brief Instrument for Assessing Healthcare
14 Employee Satisfaction in a Low-Income Setting. *PLOS ONE*, 8(11), e79053.
- 15
16 27 Chang, E. C. (2017). Measuring job satisfaction among healthcare staff in the United States: a
17 confirmatory factor analysis of the Satisfaction of Employees in health Care (SEHC) survey.
18 *International Journal for Quality in health Care*, 29(2), 262-268.
- 19
20 28 Open Data Kit. <https://opendatakit.org/>. Accessed 5th February 2021.
- 21
22 29 Pocock S, et al. Subgroup analysis, covariate adjustment and baseline comparisons in clinical trial
23 reporting: current practice and problems. *Stat Med*. 2002;21:2917-30.
- 24
25 30 The CONSORT Group. CONSORT Statement [Internet]; item 15.
- 26
27 31 R Core Team. R: A language and environment for statistical computing. R Foundation for
28 Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.
- 29
30 32 Torgerson David J. Contamination in trials: is cluster randomisation the answer? *BMJ* 2001;
31 322:355.
- 32
33 33 Benjamin-Chung J, Arnold BF, Berger D, Luby SP, Miguel E, Colford JM Jr, Hubbard AE. Spillover
34 effects in epidemiology: parameters, study designs and methodological considerations. *Int J*
35 *Epidemiol*. 2018 Feb 1;47(1):332-347. doi: 10.1093/ije/dyx201. PubMed PMID: 29106568; PubMed
36 Central PMCID: PMC5837695.
- 37
38 34 Leon N, Balakrishna Y, Hohlfeld A, Odendaal WA, Schmidt BM, Zweigenthal V, Anstey Watkins J,
39 Daniels K. Routine Health Information System (RHIS) improvements for strengthened health system
40 management. *Cochrane Database Syst Rev*. 2020 Aug 13;8:CD012012. doi:
41 10.1002/14651858.CD012012.pub2. PMID: 32803893.
- 42
43 35 Allanson ER, Tunçalp Ö, Vogel JP, Khan DN, Oladapo OT, Long Q, Gülmezoglu AM. Implementation
44 of effective practices in health facilities: a systematic review of cluster randomised trials. *BMJ Glob*
45 *Health*. 2017 Jul 20;2(2):e000266. doi: 10.1136/bmjgh-2016-000266. PMID: 29081997; PMCID:
46 PMC5656132.
- 47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3
4
5
6
7 36 Brown DW, Bosch-Capblanch X, Shimp L. Where Do We Go From Here? Defining an Agenda for
8 Home-Based Records Research and Action Considering the 2018 WHO Guidelines. *Glob Health Sci*
9 *Pract.* 2019 Mar 29;7(1):6-11. doi: 10.9745/GHSP-D-18-00431. PMID: 30877139; PMCID:
10 PM6538131.

11
12
13
14 37 Joos O, Silva R, Amouzou A, Moulton LH, Perin J, Bryce J, et al. Evaluation of a mhealth data
15 quality intervention to improve documentation of pregnancy outcomes by health surveillance
16 assistants in Malawi: a cluster randomized trial. *PloS One* 2016;11(1):e0145238. PMID: 26731401.
17 DOI: <https://doi.org/10.1371/journal.pone.0145238>

18
19
20
21 38 Langlois EV, Ranson MK, Bärnighausen T, Bosch-Capblanch X, Daniels K, El-Jardali F, Ghaffar A,
22 Grimshaw J, Haines A, Lavis JN, Lewin S, Meng Q, Oliver S, Pantoja T, Straus S, Shemilt I, Tovey D,
23 Tugwell P, Waddington H, Wilson M, Yuan B, Røttingen JA. Advancing the field of health systems
24 research synthesis. *Syst Rev.* 2015 Jul 10;4:90. doi: 10.1186/s13643-015-0080-9. PMID: 26159806;
25 PMCID: PMC4498528.

26
27
28
29 39 Bosch-Capblanch X, Lavis JN, Lewin S, Atun R, Røttingen JA, Dröschel D, Beck L, Abalos E, El-Jardali
30 F, Gilson L, Oliver S, Wyss K, Tugwell P, Kulier R, Pang T, Haines A. Guidance for evidence-informed
31 policies about health systems: rationale for and challenges of guidance development. *PLoS Med.*
32 2012;9(3):e1001185. doi: 10.1371/journal.pmed.1001185. Epub 2012 Mar 6. PMID: 22412356;
33 PMCID: PMC3295823.

34
35
36
37 40 WHO. WHO recommendations on home-based records for maternal, newborn and child health.
38 WHO 2018.

39
40
41
42 41 WHO. Recommendations on digital interventions for health system strengthening. WHO
43 guideline. WHO 2019.
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Tables and Figures

Table 1. Outcomes and parameters used to estimate the sample sizes.

	Outcome name	Subjects	Definition	Baseline estimate	Expected change	Comments
1	Vaccination adherence	Children under-1 in (sample of households in catchment areas)	Number of vaccines given in the previous calendar year over the number of vaccine due in the same period	75 given per 100 due	Increase of 10 per 100	Vaccines are clustered within children, and children within HFs
2	Data concordance	Recording tools in health facilities (samples of records)	Number of health care events (e.g. vaccinations, antenatal care consultations) recounted in the previous calendar year versus the number of health care events reported in the same time period	7 recounted for each 10 reported [3]	Increase of 1 recounted	A single estimate can be obtained in each HF or by time periods (no clustering)
3	Diagnostic scope	Records of sick child consultations (samples of records)	Number of diagnosis in each sick child consultation during the previous calendar year	30% with more than 1 diagnosis	35% with more than 1 diagnosis	Individual consultations are clustered within HF
4	Treatment appropriateness	Records of sick child consultations (samples of records)	Number of treatments correctly prescribed in each sick child consultation during the previous calendar year	Half appropriate over all consultations	Increase to 60%	Individual consultations are clustered within HF (one treatment per child)
5	Health workforce satisfaction	Health workers (all health workers form include health facilities)	Degree (score) of satisfaction across all health facilities in each arm, with the intervention	50% satisfied	75% satisfied	Maybe two or three health workers can be approached in each health facility

Figure 1. PHISICC research programme structure, processes, deliverables and flow of evidence.

Footnote to Figure 1. WS: work stream. Timelines are approximate.

1
2
3
4
5
6 **Figure 2. CONSORT diagram: trial flow chart.**
7

8
9 Separate file.
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

For peer review only

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

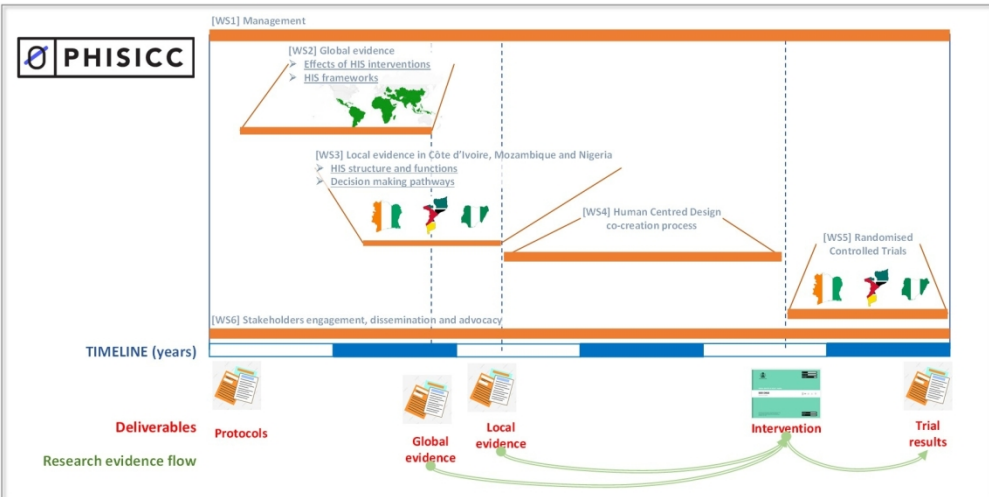


Figure 1. PHISICC research programme structure, processes, deliverables and flow of evidence.

312x157mm (150 x 150 DPI)

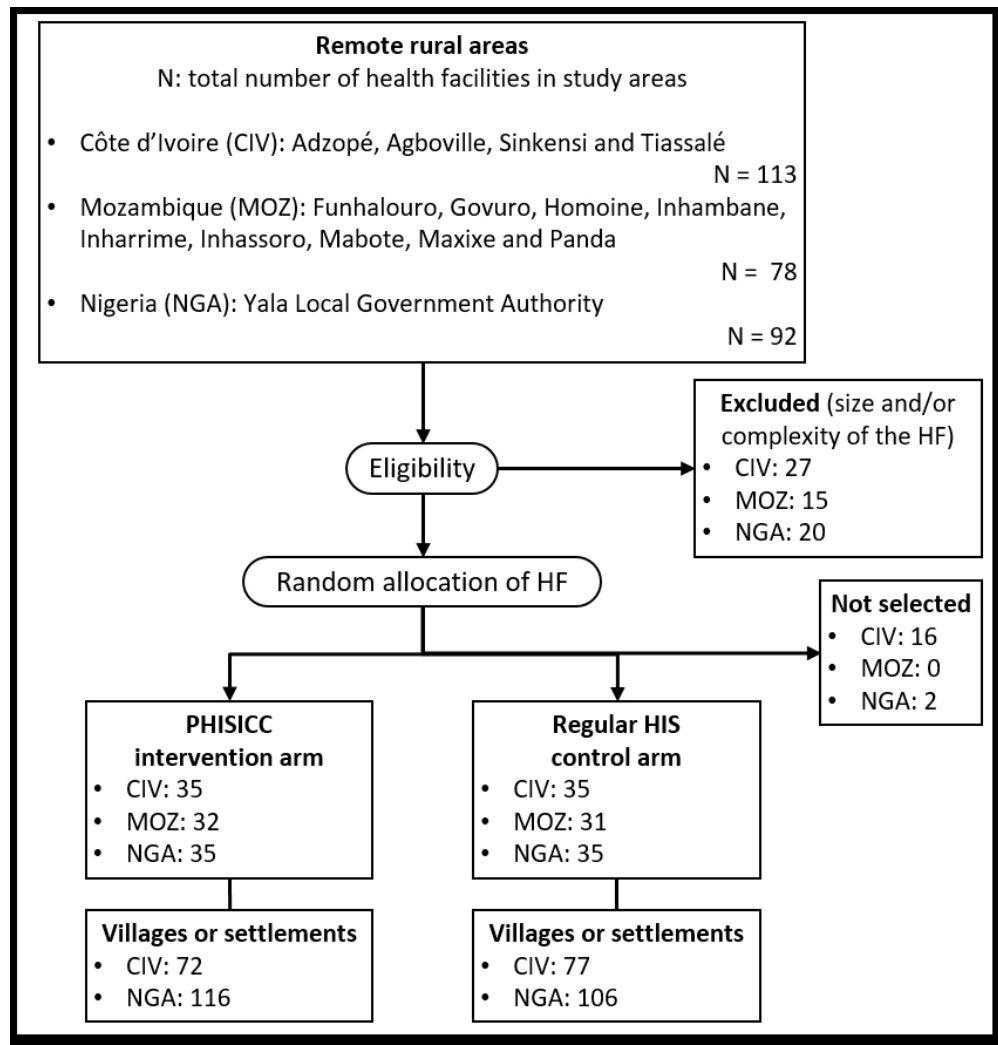


Figure 12. CONSORT diagram: trial flow chart.

144x150mm (150 x 150 DPI)

```
1
2
3 #
4 # clustersampleSize_proportions_baseline&endline.r
5 # get power of cluster randomised trial for binary outcomes (baseline and
6 # endline surveys)
7 # 2 groups (control & intervention)
8 # clustered within HF
9
10
11 rm(list=ls())
12
13 # if the package lme4 is not already installed (needed for regression with
14 # random effects)
15 # install.packages(lme4)
16 require(lme4)
17 #install.packages("reshape")
18 library(reshape)
19
20
21
22 # INPUTS
23 numGroups<-2
24 numHFPerGroup<-35
25 numTrialsToSimulate<-100
26 # numTrialsToSimulate: use 10 to test that the script runs, use 100 or 1000 for
27 # precise estimate of power
28
29
30
31
32 # choose input set and remove #s to run
33
34 # inputs for 'treatments with appropriate diagnosis'
35 pInterv<-0.60
36 pControl<-0.50
37 sdHFcluster<-0.55
38 # for k=0.1, 0.20; for k=0.25, 0.55
39 numObsPerHF<-25
40
41 # inputs for vaccination adherence
42 # proportions in interventions and control groups
43 # pInterv<-0.8
44 # pControl<-0.75
45 # sdHFcluster<-2.63
46 # numObsPerHF<-30
47
48
49
50 # inputs for 'more than one diagnosis'
51 # pInterv<-0.35
52 # pControl<-0.30
53 # sdHFcluster<-0.39
54 # for k=0.1, 0.16; for k=0.25, 0.39
55 # numObsPerHF<-60
56
57
58
59 # NB getsd is a function at the bottom of the script to turn k into sdHFcluster
60 (sdHFcluster is on the logit scale)
```

```

1
2
3
4
5
6 # --- simulation ----
7
8 # SET UP DATA STRUCTURE (intervention, HF)
9 totNumHF <- numHFPerGroup*numGroups
10 HFList<-seq(1:totNumHF)
11 interv<- rep(c(0,1),each=(totNumHF/2) )
12 intervEffect<-rep( c(0,(log(pInterv/(1-pInterv)) -
13 log(pControl/(1-pControl))) ), each=(totNumHF/2) )
14
15
16 xtemp<-cbind(interv,HFList,intervEffect)
17
18 # SET UP STORE FOR PVALUES AND PRECISION
19 storeResults<-array(-9,dim=c(numTrialsToSimulate,3))
20 colnames(storeResults)<-c("pvalue","coeff","stderr")
21
22
23 # LOOP THROUGH THE SIMULATIONS
24
25 for (i in 1:numTrialsToSimulate) {
26
27 # simulate the HF cluster effects
28 HFEffect<-rnorm(totNumHF,mean=0,sd=sdHFcluster)
29 xtemp2a<-cbind(xtemp, HFEffect)
30 xtemp2a<-data.frame(xtemp2a)
31
32
33 # get expected proportions (pre and post)
34 xtemp2a$expectedprelogodds<-log(pControl/(1-pControl)) + xtemp2a$HFEffect
35
36 xtemp2a$expectedpostlogodds<-log(pControl/(1-pControl)) +
37 xtemp2a$intervEffect + xtemp2a$HFEffect
38
39 xtemp2a$expectedpre<-exp(xtemp2a$expectedprelogodds)/(1+exp(xtemp2a$expectedpre
40 logodds))
41
42 xtemp2a$expectedpost<-exp(xtemp2a$expectedpostlogodds)/(1+exp(xtemp2a$expectedp
43 ostlogodds))
44
45 # expand by the number of observations per HF
46 xtemp2b<-untable(xtemp2a, num=numObsPerHF)
47 numObs<-dim(xtemp2b)[1]
48
49
50 # simulate individual observations from cluster mean rates
51 simObsPost<-rep(0,numObs)
52 simObsPre<-rep(0,numObs)
53 for (j in 1:numObs) {
54 simObsPost[j]<-rbinom(n=1, size=1,prob=xtemp2b$expectedpost[j])
55 simObsPre[j]<-rbinom(n=1, size=1,prob=xtemp2b$expectedpre[j])
56 }
57 # drop variables not needed further
58 xtemp2b$expectedpostlogodds<-NULL; xtemp2b$expectedprelogodds<-NULL
59
60

```

```
1
2
3
4     # stack pre and post observations
5     # get post
6     xtemp3<-cbind(xtemp2b,simObsPost)
7     xtemp3<-data.frame(xtemp3)
8     xtemp3$simObs<-xtemp3$simObsPost
9     xtemp3$simObsPost<-NULL
10    xtemp3$post<-1
11    # get pre
12    xtemp4<-cbind(xtemp2b,simObsPre)
13    xtemp4<-data.frame(xtemp4)
14    xtemp4$simObs<-xtemp4$simObsPre
15    xtemp4$simObsPre<-NULL
16    xtemp4$post<-0
17    xtemp4$interv<-0
18    xtemp5<-rbind(xtemp3,xtemp4)
19
20
21
22    # carry out analysis for individual trial
23    m <- glmer(simObs ~ as.factor(interv) + post + (1 | HFList),
24 data<-xtemp5, family=binomial)
25
26    # store result of individual trial in storeResults (p-value, coefficient
27 and std error)
28    out1<-summary(m)$coefficients
29    storeResults[i,2]<-out1[2,1]
30    storeResults[i,3]<-out1[2,2]
31    storeResults[i,1]<-out1[2,4]
32
33
34    print(i)
35
36 } # End of loop
37
38 # calculate power
39 pvalue<-storeResults[,1]
40 power<-length(pvalue[pvalue<0.05])/length(pvalue)
41
42 cat("power ", power, "\n")
43
44
45
46
47 # ----- run to here -----
48
49
50
51
52
53
54
55 # -----
56 # getsd: function to estimate between-cluster variation from k (Hayes and
57 Bennet sd/mean) and input base proportion (base0p)
58
59
60
```



```
1
2
3 getsd<-function(base0p,k){
4     sdcluster<-k*base0p
5     clusterEffect<-rnorm(1000,mean=0,sd=sdcluster)
6     expectedp<-base0p + clusterEffect
7     expectedp[expectedp>1]<-0.9999
8     expectedp[expectedp<0]<-0.0001
9     logitexpectedp<-log((expectedp)/(1-expectedp))
10    sdlog<-sd(logitexpectedp)
11    cat("estimated sdlog ", sdlog, "\n")
12
13 }
14
15 # example
16 getsd(0.30,0.25)
17
18 getsd(0.50, 0.25)
```

```

1
2
3 #
4 # clusterSampleSize_concordance.r
5 # get power of cluster randomised trial
6 # ratios (outcome is continuous)
7 # fixed to 2 groups
8 # records and reports clustered within HF
9 #
10
11
12 # if the package lme4 is not already installed (needed for regression with
13 random effects)
14 # install.packages(lme4)
15 # install.packages(lmerTest)
16 require(lme4)
17 require(lmerTest)
18
19
20
21 # EXAMPLE INPUTS
22 numGroups<-2
23 numHFPerGroup<-35
24 numReportedPerHF<-100
25 # assuming equal numbers of vaccinations per HF
26 numTrialsToSimulate<-100
27 # 100 or 1000 needed for precision of the power estimate, use 10 for test runs
28
29 ratioControl<-0.7
30 ratioInterv<-0.8
31 sdHFcluster<-0.25*0.8
32 # sdHFcluster is on the log scale, calculated using k=0.25
33
34
35
36
37 # --- run simulation from here ----
38
39 # SET UP DATA STRUCTURE (intervention, HF)
40 totNumHF<-numGroups*numHFPerGroup
41 HFList<-rep(seq(1:(numHFPerGroup*numGroups)),each=1)
42 interv<-c( rep(c(0,1),each=(totNumHF/2)))
43 intervEffect<-rep( c(0,(ratioInterv - ratioControl )), each=(totNumHF/2) )
44 xtemp<-cbind(interv,HFList,intervEffect)
45
46
47 # SET UP STORE FOR PVALUES AND PRECISION
48 storeResults<-array(-9,dim=c(numTrialsToSimulate,3))
49 colnames(storeResults)<-c("pvalue","coeff","stderr")
50
51
52 # LOOP THROUGH THE SIMULATIONS
53
54 for (i in 1:numTrialsToSimulate) {
55
56 # simulate the HF cluster effects
57
58 HFEffect<-rnorm(numHFPerGroup*numGroups,mean=0,sd=sdHFcluster)
59 xtemp2<-cbind(xtemp, HFEffect)
60

```

```

1
2
3
4     # get expected ratios (pre and post)
5     expectedpreratio<-ratioControl + HFEeffect
6     expectedpostratio<-ratioControl + intervEffect + HFEeffect
7     expectedpreratio[expectedpreratio<0.0001]<-0.0001
8     expectedpostratio[expectedpostratio<0.0001]<-0.0001
9
10    # simulate individual observations as poisson rate of number reported per
11    1 recorded
12    simObsPost<-rep(0,length(expectedpostratio))
13    simObsPre<-rep(0,length(expectedpreratio))
14    for (j in 1:length(expectedpostratio)) {
15        simObsPost[j]<-rpois(n=1,expectedpostratio[j]*numReportedPerHF)
16        simObsPre[j]<-rpois(n=1,expectedpreratio[j]*numReportedPerHF)
17    }
18
19
20
21    # stack pre and post observations
22    # post
23    xtemp3<-cbind(xtemp2,simObsPost)
24    xtemp3<-data.frame(xtemp3)
25    xtemp3$simObs<-xtemp3$simObsPost
26    xtemp3$simObsPost<-NULL
27    xtemp3$post<-1
28    # pre
29    xtemp4<-cbind(xtemp2,simObsPre)
30    xtemp4<-data.frame(xtemp4)
31    xtemp4$simObs<-xtemp4$simObsPre
32    xtemp4$simObsPre<-NULL
33    xtemp4$post<-0
34    xtemp4$interv<-0
35    # stack pre and post
36    xtemp5<-rbind(xtemp3,xtemp4)
37    xtemp5$distanceToOne<-abs(1-(xtemp5$simObs/numReportedPerHF))
38
39
40
41    # carry out analysis for individual trial
42    m <- lmer(distanceToOne ~ as.factor(interv) + post + (1|HFList),
43    data=xtemp5)
44
45    # store result of individual trial in storeResults (p-value, coefficient
46    and std error)
47    out1<-summary(m)$coefficients
48    # estimate
49    storeResults[i,2]<-out1[2,1]
50    # se
51    storeResults[i,3]<-out1[2,2]
52    # p-value
53    storeResults[i,1]<-out1[2,5]
54
55
56    print(i)
57
58    } # End of loop
59
60

```

```
1
2
3
4     # calculate power
5     pvalue<-storeResults[,1]
6     power<-length(pvalue[pvalue<0.05])/length(pvalue)
7
8     cat("power ", power, "\n")
9
10
11     # -----
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
```

Table 1: CONSORT 2010 checklist of information to include when reporting a cluster randomised trial

Section/Topic	Item No	Standard Checklist item	Extension for cluster designs	Page No *
Title and abstract				
	1a	Identification as a randomised trial in the title	Identification as a cluster randomised trial in the title	1
	1b	Structured summary of trial design, methods, results, and conclusions	See table 2	5
Introduction				
Background and objectives	2a	Scientific background and explanation of rationale	Rationale for using a cluster design	4
	2b	Specific objectives or hypotheses	Whether objectives pertain to the cluster level, the individual participant level or both	5
Methods				
Trial design	3a	Description of trial design (such as parallel, factorial) including allocation ratio	Definition of cluster and description of how the design features apply to the clusters	6
	3b	Important changes to methods after trial commencement (such as eligibility criteria), with reasons		Not applicable
Participants	4a	Eligibility criteria for participants	Eligibility criteria for clusters	6
	4b	Settings and locations where the data were collected		6
Interventions	5	The interventions for each group with sufficient details to allow replication, including how and when they were actually administered	Whether interventions pertain to the cluster level, the individual participant level or both	8
Outcomes	6a	Completely defined pre-specified primary and secondary outcome measures, including how and when they were assessed	Whether outcome measures pertain to the cluster level, the individual participant level or both	10 and Table 1
	6b	Any changes to trial outcomes after the trial commenced, with reasons		Not applicable
Sample size	7a	How sample size was determined	Method of calculation, number of clusters(s) (and whether equal or unequal cluster sizes are assumed), cluster size, a coefficient of intracluster correlation (ICC or <i>k</i>), and an indication of its uncertainty	12
	7b	When applicable, explanation of any interim analyses and stopping guidelines		Not yet applicable

Section/Topic	Item No	Standard Checklist item	Extension for cluster designs	Page No *
Randomisation:				
Sequence generation	8a	Method used to generate the random allocation sequence		7, 8
	8b	Type of randomisation; details of any restriction (such as blocking and block size)	Details of stratification or matching if used	Not applicable
Allocation concealment mechanism	9	Mechanism used to implement the random allocation sequence (such as sequentially numbered containers), describing any steps taken to conceal the sequence until interventions were assigned	Specification that allocation was based on clusters rather than individuals and whether allocation concealment (if any) was at the cluster level, the individual participant level or both	7
Implementation	10	Who generated the random allocation sequence, who enrolled participants, and who assigned participants to interventions	Replace by 10a, 10b and 10c	7
	10a		Who generated the random allocation sequence, who enrolled clusters, and who assigned clusters to interventions	7
	10b		Mechanism by which individual participants were included in clusters for the purposes of the trial (such as complete enumeration, random sampling)	10
	10c		From whom consent was sought (representatives of the cluster, or individual cluster members, or both), and whether consent was sought before or after randomisation	8
Blinding	11a	If done, who was blinded after assignment to interventions (for example, participants, care providers, those assessing outcomes) and how		8
	11b	If relevant, description of the similarity of interventions		Not applicable
Statistical methods	12a	Statistical methods used to compare groups for primary and secondary outcomes	How clustering was taken into account	15
	12b	Methods for additional analyses, such as subgroup analyses and adjusted analyses		15

Section/Topic	Item No	Standard Checklist item	Extension for cluster designs	Page No *
Results				Not yet applicable (protocol manuscript)
Discussion				
Limitations	20	Trial limitations, addressing sources of potential bias, imprecision, and, if relevant, multiplicity of analyses		18
Generalisability	21	Generalisability (external validity, applicability) of the trial findings	Generalisability to clusters and/or individual participants (as relevant)	18,19
Interpretation	22	Interpretation consistent with results, balancing benefits and harms, and considering other relevant evidence		Not yet applicable
Other information				
Registration	23	Registration number and name of trial registry		1
Protocol	24	Where the full trial protocol can be accessed, if available		1
Funding	25	Sources of funding and other support (such as supply of drugs), role of funders		20

* Page numbers: as seen in the document "draft_Proof_hi.pdf" (which has 34 pages)

CONSORT 2010 checklist of information to include when reporting a cluster randomised trial

Section/Topic	Item No	Standard Checklist item	Extension for cluster designs	Page No *
Title and abstract				
	1a	Identification as a randomised trial in the title	Identification as a cluster randomised trial in the title	1
	1b	Structured summary of trial design, methods, results, and conclusions	See table 2	5
Introduction				
Background and objectives	2a	Scientific background and explanation of rationale	Rationale for using a cluster design	4
	2b	Specific objectives or hypotheses	Whether objectives pertain to the cluster level, the individual participant level or both	5
Methods				
Trial design	3a	Description of trial design (such as parallel, factorial) including allocation ratio	Definition of cluster and description of how the design features apply to the clusters	6
	3b	Important changes to methods after trial commencement (such as eligibility criteria), with reasons		Not applicable
Participants	4a	Eligibility criteria for participants	Eligibility criteria for clusters	6
	4b	Settings and locations where the data were collected		6
Interventions	5	The interventions for each group with sufficient details to allow replication, including how and when they were actually administered	Whether interventions pertain to the cluster level, the individual participant level or both	8
Outcomes	6a	Completely defined pre-specified primary and secondary outcome measures, including how and when they were assessed	Whether outcome measures pertain to the cluster level, the individual participant level or both	10 and Table 1
	6b	Any changes to trial outcomes after the trial commenced, with reasons		Not applicable
Sample size	7a	How sample size was determined	Method of calculation, number of clusters(s) (and whether equal or unequal cluster sizes are assumed), cluster size, a coefficient of intracluster correlation (ICC or k), and an indication of its uncertainty	12
	7b	When applicable, explanation of any interim analyses and stopping guidelines		Not yet applicable

Section/Topic	Item No	Standard Checklist item	Extension for cluster designs	Page No *
Randomisation:				
Sequence generation	8a	Method used to generate the random allocation sequence		7, 8
	8b	Type of randomisation; details of any restriction (such as blocking and block size)	Details of stratification or matching if used	Not applicable
Allocation concealment mechanism	9	Mechanism used to implement the random allocation sequence (such as sequentially numbered containers), describing any steps taken to conceal the sequence until interventions were assigned	Specification that allocation was based on clusters rather than individuals and whether allocation concealment (if any) was at the cluster level, the individual participant level or both	7
Implementation	10	Who generated the random allocation sequence, who enrolled participants, and who assigned participants to interventions	Replace by 10a, 10b and 10c	7
	10a		Who generated the random allocation sequence, who enrolled clusters, and who assigned clusters to interventions	7
	10b		Mechanism by which individual participants were included in clusters for the purposes of the trial (such as complete enumeration, random sampling)	10
	10c		From whom consent was sought (representatives of the cluster, or individual cluster members, or both), and whether consent was sought before or after randomisation	8
Blinding	11a	If done, who was blinded after assignment to interventions (for example, participants, care providers, those assessing outcomes) and how		8
	11b	If relevant, description of the similarity of interventions		Not applicable
Statistical methods	12a	Statistical methods used to compare groups for primary and secondary outcomes	How clustering was taken into account	15
	12b	Methods for additional analyses, such as subgroup analyses and adjusted analyses		15

Section/Topic	Item No	Standard Checklist item	Extension for cluster designs	Page No *
Results				Not yet applicable (protocol manuscript)
Discussion				
Limitations	20	Trial limitations, addressing sources of potential bias, imprecision, and, if relevant, multiplicity of analyses		18
Generalisability	21	Generalisability (external validity, applicability) of the trial findings	Generalisability to clusters and/or individual participants (as relevant)	18,19
Interpretation	22	Interpretation consistent with results, balancing benefits and harms, and considering other relevant evidence		Not yet applicable
Other information				
Registration	23	Registration number and name of trial registry		1
Protocol	24	Where the full trial protocol can be accessed, if available		1
Funding	25	Sources of funding and other support (such as supply of drugs), role of funders		20

BMJ Open

Does an innovative paper-based health information system (PHISICC) improve data quality and use in primary health care? Protocol of a multi-country, cluster randomised controlled trial in Sub-Saharan African rural settings.

Journal:	<i>BMJ Open</i>
Manuscript ID	bmjopen-2021-051823.R2
Article Type:	Protocol
Date Submitted by the Author:	02-Jul-2021
Complete List of Authors:	Bosch-Capblanch, Xavier; Swiss Tropical and Public Health Institute; University of Basel Oyo-Ita, Angela; University of Calabar, Department of Community Medicine Muloliwa, Artur; Lurio University Faculty of Health Sciences Yapi, Richard; Centre Suisse de Recherches Scientifiques en Côte d'Ivoire Auer, Christian ; Swiss Tropical and Public Health Institute; University of Basel Samba, Mamadou; Ministère de la Santé et de l'Hygiène Publique Gajewski, Suzanne; Swiss Tropical and Public Health Institute; University of Basel Ross, Amanda; Swiss Tropical and Public Health Institute; University of Basel Krause, L Kendall; Bill & Melinda Gates Foundation Ekpenyong, Nnette; University of Calabar Teaching Hospital, Department of Community Medicine Nwankwo, Ogonna; University of Calabar, Department of Community Medicine Njepuome, Anthonia; Swiss Tropical and Public Health Institute Lee, Sofia; Swiss Tropical and Public Health Institute Sacarlal, Jahit ; Eduardo Mondlane University, Faculty of Medicine Madede, Tavares; Universidade Eduardo Mondlane Berté, Salimata; Centre Suisse de Recherches Scientifiques en Côte d'Ivoire Matsinhe, Graça; Ministério da Saúde Garba, Abdullahi; National Primary Healthcare Development Agency Brown, David; BCGI LLC / pivot-23.5°
Primary Subject Heading:	Health services research
Secondary Subject Heading:	Global health, Health policy, Research methods
Keywords:	Human resource management < HEALTH SERVICES ADMINISTRATION & MANAGEMENT, Organisation of health services < HEALTH SERVICES ADMINISTRATION & MANAGEMENT, Protocols & guidelines < HEALTH SERVICES ADMINISTRATION & MANAGEMENT, Quality in health care <

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

	HEALTH SERVICES ADMINISTRATION & MANAGEMENT, STATISTICS & RESEARCH METHODS, PRIMARY CARE

SCHOLARONE™
Manuscripts



I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our [licence](#).

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which [Creative Commons](#) licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

Title

Does an innovative paper-based health information system (PHISICC) improve data quality and use in primary health care? Protocol of a multi-country, cluster randomised controlled trial in Sub-Saharan African rural settings.

Authors

- Xavier Bosch-Capblanch. Corresponding author. Swiss Tropical and Public Health Institute. Basel, Switzerland. And University of Basel, Basel, Switzerland. x.bosch@unibas.ch
- Angela Oyo-Ita. Department of Community Medicine, University of Calabar, Calabar, Nigeria. oyo_ita@yahoo.com
- Artur Manuel Muloliwa. Faculty of Health Sciences, Lúrio University, Nampula, Mozambique. muloliwa@yahoo.com.br
- Richard B Yapi. Centre Suisse de Recherches Scientifiques en Côte d'Ivoire, Abidjan, Côte d'Ivoire. And Centre d'Entomologie Médicale et Vétérinaire, Université Alassane Ouattara, Bouaké, Côte d'Ivoire. richard.yapi@csrs.ci
- Christian Auer. Swiss Tropical and Public Health Institute. Basel, Switzerland. And University of Basel, Basel, Switzerland. christian.auer@swisstph.ch
- Mamadou Samba. Ministère de la Santé et de l'Hygiène Publique, Abidjan, Côte d'Ivoire. And Université Félix Houphouët Boigny, Abidjan, Côte d'Ivoire. samba.mamadou@gmail.com
- Suzanne Gajewski. Swiss Tropical and Public Health Institute. Basel, Switzerland. And University of Basel, Basel, Switzerland. suzanne.gajewski@swisstph.ch
- Amanda Ross. Swiss Tropical and Public Health Institute. Basel, Switzerland. And University of Basel, Basel, Switzerland. amanda.ross@swisstph.ch
- L Kendall Krause. Bill & Melinda Gates Foundation, Seattle, USA. Kendall.Krause@gatesfoundation.org
- Nnette Ekpenyong. Department of Community Medicine, University of Calabar Teaching Hospital, Calabar, Nigeria. And Department of Community Medicine University of Calabar, Calabar, Nigeria. nnekon2015@gmail.com
- Ogonna Nwankwo. Swiss Tropical and Public Health Institute, Basel, Switzerland. And University of Basel, Basel, Switzerland. And Health Policy, Management and Economics Firm, Department of Community. Medicine, University of Calabar, Calabar, Nigeria. ogonna.nwankwo@swisstph.ch
- Anthonia Ngozi Njebuome. Swiss Tropical and Public Health Institute, Abuja, Nigeria. ngonjep@yahoo.com
- Sofia Mandjate Lee. Swiss Tropical and Public Health Institute; Manhiça, Mozambique. sofia_mandjate@yahoo.com.br
- Jahit Sacarlal. Faculty of Medicine, Eduardo Mondlane University, Maputo, Mozambique. jahityash2002@gmail.com
- Tavares Madede. Department of Community Health, Faculty of Medicine, Eduardo Mondlane University, Maputo, Mozambique. tmadede@gmail.com
- Salimata Berté. Centre Suisse de Recherches Scientifiques en Côte d'Ivoire, Abidjan, Côte d'Ivoire. sali.berte@csrs.ci
- Graça Matsinhe. Expanded Program on Immunization, Ministry of Health, Maputo, Mozambique. gmatsinhe@gmail.com.
- Abdullahi Bulama Garba. Director Planning, Research and Statistics, National Primary Health Care Development Agency, Abuja, Nigeria. ab.garba@nphcda.gov.ng.
- David W Brown. Managing Principal, BCGI LLC / pivot-23.5°, North Carolina, USA. david.brown@pivot235.org

Abstract

Introduction

Frontline health workers in remote health facilities are the first contact of the formal health sector and are confronted with life-saving decisions. Health information systems (HIS) support the collection and use of data. However, HIS focus on reporting and are unfit to support decisions. Since data tools are paper-based in most primary health care settings, we have produced an innovative paper-based HIS (PHISICC) using a Human Centred Design approach. We are carrying out a cluster-randomised controlled trial in three African countries to assess the effects of PHISICC compared with the current systems.

Methods and analysis

Study areas are in rural zones of Côte d'Ivoire, Mozambique and Nigeria. Seventy health facilities in each country have been randomly allocated to using PHISICC tools or to continuing to use the regular HIS tools. We have randomly selected households in the catchment areas of each health facility to collect outcomes' data. The baseline survey has been carried out in two of the three countries, the end-line survey is planned for mid-2021. Primary outcomes include data quality and use and coverage of health services and health workers satisfaction; secondary outcomes are additional data quality and use parameters, childhood mortality and additional health workers and clients experience with the system. Just prior to the implementation of the trial we had to relocate the studies in Mozambique and Côte d'Ivoire due to unforeseen logistical issues. The effects of the intervention will be estimated using regression models and accounting for clustering using random effects.

Ethics and dissemination

Ethics committees in Côte d'Ivoire, Mozambique and Nigeria approved the trials. We plan to disseminate our findings, data and research materials among researchers and policy makers. We aim at having our findings included in systematic reviews on health systems interventions and future guidance development on the matter.

Registration

Pan African Clinical Trials Registry - PACTR201904664660639. Registered 01/04/2019,
<https://pactr.samrc.ac.za/Search.aspx>.

Article summary

Strengths and limitations of the study

- These trials assess the effects of improving paper-based health information systems, which are greatly used particularly in remote, rural areas but which are neglected in research.
- The paper-based interventions have been developed using a Human Centred Design approach, with frontline health workers and designers driving the co-creation process.
- Despite the complexity of health systems interventions, we have applied robust experimental methods, together with qualitative research, to assess and understand the effects of the paper-based intervention. Robust evidence on health systems is more likely to gain the credibility of policy-makers and to make it into systematic reviews, guidance development and policy and practice.
- Research targeting frontline health workers in remote, rural areas has to take place where they live and work, which poses serious obstacles in the organisation, management and monitoring of the trials.
- These obstacles, aggravated by the COVID-19 pandemic, have challenged the mobility of the research team, the availability of the intervention in one of the countries and the duration of the trials.

Introduction

Frontline health workers (HW) in remote, rural health facilities (HF) in many countries are the first contact with the formal health sector of the population and they are confronted with life-saving clinical and public health decisions on a daily basis. Decisions are made by exerting a balanced judgment on the information related to health care events, such as making the correct diagnoses or deciding on which vaccinations a child should receive on a given day. In order to properly handle this information, appropriate data support tools and processes are required, referred to as the health information system (HIS); or Routine HIS or Health Management Information System [1]. In reality, though, HIS are primarily designed to report aggregated health events to the higher tiers of the health systems rather than to inform decision-making at the point of care [2].

Increasing pressure by donors and governments to collect more and more data has aggravated the situation, through the proliferation of data support tools that have overloaded frontline health workers compromising their capacity to deliver good quality of care and to delivery good quality data [3], for higher level decision-making.

Promising 'quick fixes', such as the scale up of digital HIS, are taking a long time to implement and face enormous challenges related to infrastructure, equipment and services necessary to run them. Besides, research evidence on the effects of digital solutions remains patchy and inconsistent, even in high-income country settings, where complaints about computerisation of clinical care have been raised [4,5]. Hence, it is very likely that paper tools will remain a primary, if not unique, data support mechanism particularly in remote, rural HF in many countries.

PHISICC (Paper-based Health Information System in Comprehensive Care) is a multi-year, multi-country, mixed-methods research project that aims at producing and testing an innovative paper-based HIS to improve data quality and use, decision making and health outcomes, at Primary Health Care (PHC). It is being carried out in selected areas within Côte d'Ivoire, Mozambique and Nigeria. The project started in 2015, producing a systematic review on the effects of HIS interventions [6,7]

1
2
3 and a framework synthesis on how HIS are understood in the literature in order to learn from past
4 experiences in HIS developments. This global evidence was coupled with studies to characterise the
5 existing HIS in the three countries, to understand how health workers interact with the HIS and to
6 identify entry points for HIS design improvements. With these bodies the research team was well
7 equipped to engage into a Human Centred Design (HCD) co-creative process with frontline HW to
8 design an innovative HIS (PHISICC). See Figure 1 for an illustration of the structure, processes and
9 evidence flow within PHISICC.

10
11
12
13
14
15
16
17
18
19 The impact of the PHISICC HIS on data quality and use, quality of health care and HW perceptions is
20 being assessed concurrently in rural areas in the three countries. We describe the design of the trial
21 here, consistent with CONSORT reporting guidelines [8] and the extension for cluster randomised
22 controlled trials (CRCT) [9]; see Additional file 1.

23 24 25 26 27 28 29 **Methods**

30 31 32 **Aim**

33
34
35 The aim of the trial is to address the research question: what are the effects of an innovative paper-
36 based HIS (PHISICC) on data use and quality, quality of health and HW perceptions compared with
37 the current HIS, in rural PHC settings?
38
39
40
41
42

43 **Patient and public involvement**

44
45 There was no public or patient involvement in the design of the study or selection of study areas
46 because the intervention being assessed in these trials target health care providers and decision-
47 makers, rather than patients or the public in general. Population in the catchment area of selected
48 health facilities, potentially using their services, were only approached in order to assess the studies
49 outcomes.
50
51
52
53
54
55

56 On the other hand, we have involved health systems stakeholders and frontline health workers.

57
58
59 Ministries of Health at several levels participated in the preparation of the research proposal
60

1
2
3 (personal consultations), in the characterisation of health information systems that preceded the
4 trials (countries workshops), and throughout all project components (additional workshops,
5 newsletters and personal communication). Frontline health workers in the three countries have co-
6 created the intervention (i.e. paper based tools) through workshops, personal feedback and piloting
7 under real live conditions. Some of them are part of the research team and co-authoring this
8 manuscript.
9
10
11
12
13
14
15
16

17 **Study design**

18
19
20 The study is a CRCT in each of the three countries. In each setting, 70 health facilities are randomised
21 to intervention or control (35 per arm). The intervention arm HF use the new PHISICC tools
22 (substituting the usual HIS tools) and the control arm HF use the regular HIS tools. The trial is
23 implemented in the real life contexts of HF carrying out their usual duties.
24
25
26
27
28

29 The trials started between the end of 2019 and beginning of 2020, depending on the country, when
30 the intervention was installed and the baseline surveys carried out. Data collection will last until mid-
31 2021.
32
33
34
35
36

37 **Study areas**

38
39 Ministries of Health (MOH) officials in several countries were contacted before submitting the
40 proposal to the funding agency in order to explore the willingness to engage in a project focusing on
41 paper-based tools. Officials in several countries rejected the offer on the grounds of upcoming
42 digitalisation plans of the HIS in the country. We partnered with MOH that found the research
43 relevant to their context in three countries.
44
45
46
47
48
49

50
51 In each country, the eligibility criteria of study areas were that they had to belong to the operational
52 area of research partners; contain a large enough number of health facilities and their catchment
53 population; include vulnerable population (e.g. with low vaccination coverage, high childhood
54 mortality); and be comparatively neglected in terms of infrastructure and services. We excluded
55 areas with concurrent research or other types of activities that could conflict with the CRCT (such as
56
57
58
59
60

1
2
3 the co-existence of another health-related study, massive developments in infrastructure or activities
4 involving migration of the population, such as temporary work sites or changes in working sites) and
5 areas with threats to safety or security that could jeopardise research activities.
6
7
8

9
10 The study areas are located in Adzopé, Agboville, Tiassalé and Sikensi districts (Côte d'Ivoire); in
11 Funhalouro, Govuro, Homoine, Inhambane, Inharrime, Inhassoro, Mabote, Maxixe and Panda
12 (Inhambane province, Mozambique); and in Yala Local Government Authority (Cross River State,
13
14
15
16
17
18 Nigeria).

19 **Eligibility of health facilities**

20
21 The intervention is implemented at the HF level. The eligibility criteria of the HF were that they had
22 to be located in the study areas, belong to the governmental health sector and their main activity
23 should be the delivery of PHC services. HF were excluded if they had specialised clinical services,
24 inpatients, physicians providing care or with plans for staff turn-over involving intervention and
25 control HF.
26
27
28
29
30
31

32
33 A 'master list' of eligible health facilities was prepared based on information provided by the MOH
34 across all study areas. We aimed at selecting 70 of the eligible HF in each country, using simple
35 random techniques in R [10]. See in Figure 2 the selection and allocation trial flow chart.
36
37
38
39

40 **Allocation and blinding**

41
42 Allocation of the 70 HF per country into the intervention and control arms took place in a formal
43 event, gathering research partners and MOH officials to offer transparency and promote study
44 ownership by local and national authorities. Equally sized, folded pieces of paper with the names and
45 codes of included HF written on them were introduced in an opaque receptacle where they were
46 manually and blindly mixed. A second receptacle contained two equally sized pieces of paper, one
47 with the word 'intervention' and another one with the word 'control'. A selected person in the
48 meeting, not belonging to the research team, extracted one piece of paper at a time to reach half the
49 number of included HF. Then, a paper was extracted from the second receptacle to assign those HF
50
51
52
53
54
55
56
57
58
59
60

1
2
3 to the intervention or control arms. The rest of the papers were extracted as well to verify
4
5 completeness and no duplication of names, and those HF assigned to the other arm.
6

7
8 Once HF were selected, all villages or settlements for each health facility catchment area were listed
9
10 and three in each catchment area were selected. In practice, we selected all villages because the
11
12 numbers were below (in Côte d'Ivoire) or just above (in Nigeria) the needs. For each village, we used
13
14 Google satellite maps to identify and geo-locate every visible roof. Where there were many houses
15
16 per village (roughly, more than fifty or so), a researcher would mark four points in the map slightly
17
18 beyond the northernmost, southernmost, easternmost and westernmost roofs seen and 30 random
19
20 points were selected within that square. From the mapped points, 10 per village (with 10 more acting
21
22 as reserve) were randomly selected and marked on another map used in the field for data collectors
23
24 to approach households. Where technical problems impeded this approach in a given village, a field
25
26 supervisor would rotate a bottle on the floor towards the centre of the village and would select at
27
28 random 10 households in the direction pointed by the bottle, from the outer limit of the village till
29
30 the centre [11].
31
32
33

34
35 Blinding is only feasible for the research team members carrying out the CRCT data collection and the
36
37 analyses of the CRCT findings. The intervention (i.e. paper tools) are by design very different from the
38
39 existing system and it is not possible to blind participants or principal researchers.
40

41
42 We already had the agreement of the MOH and selected HF compliant with the inclusion criteria
43
44 were provided with the intervention shortly after completing the baseline data collection. Therefore,
45
46 recruitment as such took place at the same time of the allocation of HF into intervention and control
47
48 arms.
49

50 51 **The intervention**

52
53 The PHISICC paper-based intervention is a full set of paper-based tools to support decision-making by
54
55 frontline HW. These are the only tools to be used by HW in the intervention arm. The PHISICC tools
56
57 encompass the whole system (i.e. recording and reporting) and all clinical and public health care
58
59
60

1
2
3 areas and are characterised by a common visual language (e.g. spaces for digits and text), and
4
5 standardised formats across health care areas. To support frontline health workers decision-making,
6
7 the PHISICC tools incorporate specific places to explicitly record critical data items (e.g. respiratory
8
9 rate in infants), graphic artefacts to distinguish severity degrees of signs or symptoms (i.e. small
10
11 square for 'normality', diamond for 'attention' and bold diamond for 'critical severity'); space to
12
13 document diagnoses and treatment decisions; and aides memoires in the first page of register
14
15 books..
16
17

18
19 The PHISICC tools have been developed from May 2017 till June 2019, including production, using a
20
21 Human Centred Design approach [12]. A strength of the Human Centred Design approach is its ability
22
23 to unlock the user's perspective so that designers can build solutions that are fully reality-based and
24
25 work well. Co-creation groups were formed in each country with researchers, staff from partner
26
27 institutions and healthcare workers, led by a team of professional designers. Research team
28
29 members supervised and coordinated exclusively the feedback on the contents of the tools, to
30
31 ensure compliance with each country clinical guidelines. At the outset of the process, the design
32
33 focused on three health care areas (i.e. antenatal care, vaccination and sick child) and slowly
34
35 extended the new visual language to other health care areas. Initial workshops served to brainstorm
36
37 on problems and potential design solutions, without any other rule than being comprehensive and
38
39 not rejecting a single idea. Designers, then, formalised some of the most promising solutions and a
40
41 first round of exchanges within the co-creation team was used to address misinterpretations or
42
43 inconsistencies. There were two in-the-field testing rounds in Mozambique, two in Côte d'Ivoire and
44
45 three in Nigeria and uncountable exchanges through teleconferences and email, in-between. The
46
47 prototypes were considered final when no errors were detected, were compliant with data needs in
48
49 each country and comments from the field could not be accommodated in the design concept or
50
51 there was no consensus on minor or formal issues being raised.
52
53
54
55

56
57 The PHISICC tools have been produced in French for Côte d'Ivoire, in Portuguese for Mozambique
58
59 and in English for Nigeria, which are the official languages used in the health systems in the three
60

1
2
3 countries. They include the official logo of the MOHs. Health care areas covered include: family
4
5 planning, antenatal care, including tetanus toxoid vaccination, delivery, post-natal care, vaccination,
6
7 sick child, adults outpatient consultation, tuberculosis diagnosis and treatment, and HIV. Referral
8
9 forms were also designed.

10
11
12 The PHISICC tools have three sub-components: registers, tallies and reports. Registers are formed by
13
14 seven DIN-A3 and one DIN-A4 (for referrals) book covering all health care areas except for
15
16 tuberculosis treatment, for which DIN-A3 cards were used. Register books have 100, 200 or 400
17
18 pages depending on the country and health care area. They are used to record individual clients' data
19
20 for each health care event, either of clinical or public health nature. Some register books have clinical
21
22 notes at the very beginning, as 'aide memoires', and an example of a filled-in form, to assist HW
23
24 when doubting how to proceed.
25
26

27
28 Tallies are DIN-A3 single sheets which contain a list of the indicators to be transferred to higher levels
29
30 of the health system, with a series of small ovals, grouped in fives, to mark with tally sticks with a
31
32 pen. In contrast to the current systems that have no tallies or only for vaccination, tallies were
33
34 created for all health care areas. In the middle-right side of the tally, a column accommodates cells
35
36 aligned with the ovals to insert the count for each indicator; and in the far right of the sheet there is
37
38 a replica of the count column, separated with a perforated line, which is detached and sent, as part
39
40 of the monthly report to the higher level in the health system.
41
42

43
44 While current HIS tools are consistently organised in tabular formats and books, where each clinical
45
46 event is recorded in a row and each variable (e.g. age, gender, HIV status, diagnosis) in a column,
47
48 PHISICC tools incorporated several innovations; in summary: a visual language to guide the clinical
49
50 decisions of health workers based on severity (i.e. if it is recorded that a child has convulsions, a
51
52 visual artefact indicates severity), more space for clinical data (e.g. vital signs), inclusion of all critical
53
54 information to assess patients (e.g. obstetric history, gestational age, fundus height, breath rate in
55
56 infants), consolidation of information of all antenatal care visits for a single pregnant woman in the
57
58
59
60

1
2
3 same page, among many other formal and contents improvements, including improved aesthetics. A
4
5 systematic comparison of the new (intervention) and current (control) tools is provided in Table 1.
6

7
8 We aimed at creating “a system” (not just some tools) focusing on decision making by frontline
9
10 health workers. The epidemiological and public health contexts in the three countries are similar, as
11
12 confirmed by the similarities in the existing HIS between the three. The visual language and the
13
14 recording forms were common to the three countries because clinical decisions are common to the
15
16 three; although forms allowed for specific tests or treatments. The reporting component was
17
18 adapted to each country set of indicators, although the visual language and reporting processes were
19
20 harmonised.
21
22

23
24 During three or four days, HW were trained on HIS before the start of the trial. In the intervention
25
26 arm they were trained on the PHISICC tools; and the control arm received a refresher training about
27
28 the regular tools, during the same number of days.
29

30
31 Additionally, given that the regular tools already contained information on past vaccination history of
32
33 children still to complete their vaccination schedule, we created a mechanism to retrieve data of
34
35 children’s vaccination status to transcribe into the new vaccination register book in the intervention
36
37 arm (‘system transition’).
38

39
40 Tools were endorsed by MOH, printed in local printing companies and distributed to HW at the end
41
42 of the training sessions. A digital spreadsheet was created to monitor consumption and order
43
44 additional tools to cover health facility needs during the life of the trial.
45
46

47 **Outcomes**

48
49
50
51 **There are five primary outcomes (**
52
53
54
55
56
57
58
59
60

1
2
3 Table 2). Vaccination adherence is defined as the total number of vaccine doses given in the correct
4
5 time interval to children in households in the health facilities catchment villages of those over the
6
7 total number of vaccine doses that should have been given during the same period. Antenatal care
8
9 visits uptake will also be considered depending on the expected number of pregnancies in the study
10
11 areas. Both are used as proxies for health outcomes in terms of protection against disease [13] and
12
13 prevention of pregnancy complications [14] and are assessed in a random sample of households in
14
15 the health facilities catchment areas. Data concordance is defined as the level of agreement of HIS
16
17 indicators between (i) records of health care events (re-counts), (ii) tallies (re-counts) and (iii) reports
18
19 (aggregated data to higher levels of the system) [3]. Decision-making will be assessed considering the
20
21 diagnostics scope in the sick child (i.e. number of different diagnoses per sick child consultation) and
22
23 treatment appropriateness (i.e. number of prescribed treatments that are supported by a
24
25 documented diagnosis). These outcomes will be assessed in a random sample of records and
26
27 corresponding reports during the last four months of the study period. Health workers satisfaction
28
29 will be assessed in all health workers in included health facilities using a standardised questionnaire
30
31 [15,16,17]. While the intervention targets HF, some of the outcomes are measured at the level of
32
33 HF, and some from patients clustered within HF catchment areas.

34
35
36
37
38
39 Secondary outcomes are classified under the following domains: data quality, data use, mortality,
40
41 HW experience, client experience and resource consumption:

- 42
43 • Data quality, assessed in a sample of records
 - 44
45 ○ Completeness of recording and reporting in specific forms; i.e. prevalence of unduly
46
47 missing data items; partograph used;
 - 48
49 ○ accuracy of recorded figures in comparison to real events (e.g. physical counting of
50
51 commodities, such as number of 500mg Paracetamol tablets as recorded versus
52
53 number of 500mg Paracetamol tablets as counted);
 - 54
55 ○ timeliness of reporting, as documented by time stamps in forms;
 - 56
57 ○ loss of data or data which does not reach the next upper administrative level.
58
59
60

- 1
- 2
- 3
- 4
- 5
- 6
- 7
- 8
- 9
- 10
- 11
- 12
- 13
- 14
- 15
- 16
- 17
- 18
- 19
- 20
- 21
- 22
- 23
- 24
- 25
- 26
- 27
- 28
- 29
- 30
- 31
- 32
- 33
- 34
- 35
- 36
- 37
- 38
- 39
- 40
- 41
- 42
- 43
- 44
- 45
- 46
- 47
- 48
- 49
- 50
- 51
- 52
- 53
- 54
- 55
- 56
- 57
- 58
- 59
- 60
- Data use, assessed in a sample of records
 - in terms of knowledge (e.g. vaccines due based on date of birth; weight for length assessments);
 - cases of different conditions properly treated in (e.g. diarrhoea cases given oral rehydration therapy according to national guidelines; pneumonia cases given appropriate antibiotic according to national guidelines;
 - public health decisions: availability of lost to follow up lists or plans for vaccination, tuberculosis and or HIV/AIDS treatment control;
 - occurrence of stock outs of essential drugs.
- Overall under-5s mortality and under-5s mortality excluding peri-natal mortality [18], in a sample of households in health facilities catchment areas.
- Health workers' 'human experience' and satisfaction (all health workers).
- District health information officers' 'human experience' (selected health care programme managers).
- Clients' 'human experience' and satisfaction, in a sample of households in health facilities catchment areas.
- Resources consumption (e.g. time use, costs)
 - intervention costs: tools, training, start-up;
 - time used for recording and reporting (e.g. time-motion study) [19];
 - cost-effectiveness per unit of additional improvement in outcomes of interest.

It is worthwhile to note that outcomes that do not relate to data quality and use will be assessed using additional data collection tools (e.g. survey questionnaires), which are the same for intervention and control health facilities. Hence, the effects of the intervention cannot be attributed to the changes in performance of the paper tools routinely used to record health care events in intervention and control health facilities, which are different by design.

1
2
3 In addition, we will consider ‘explanatory outcomes’ that will help to understand how the measured
4 effects have taken place and why. We will look at the details of the interplay between the
5 intervention, the system, the users and the context. Process indicators will be based on the
6 documented activities that have taken place, from the conception of the intervention, up to its
7 implementation, monitoring and evaluation. Process indicators may include: intervention set up and
8 implementation, monitoring of the use of the intervention, special activities targeted at vulnerable
9 populations, district reactions related to the intervention, handling of data coming from the new
10 system, sustainability based on costs information and perceptions, alignment with national health
11 policies and donor priorities. We will also explore health care services characteristics looking at
12 generic indicators from health facilities, such human resources profiles and relations with the
13 communities, population characteristics and system and context characteristics captured in early
14 stages of the project, where data are available.

31 **Sample size calculations**

32
33 The required sample sizes for each primary outcome were determined using simulation to
34 incorporate the clustering easily (Table 1) and to take the baseline and end-line surveys into account.
35 Briefly, we simulated 1000 trials with variation between them caused by drawing different samples
36 from the same distributions. We then used the regression models detailed in the data section to
37 analyse each of the simulated trials and estimate the power as the proportion of trials which
38 detected the effect of the intervention as significant. The simulation code was written in R
39 (supplementary files 1 and 2).

40
41 For each country, we require the probability of α , a type I error (rejecting the null hypothesis when it
42 is actually true) to be less 0.05 and the power to be at least 80%.

43
44 For vaccination adherence, using a sample size of 35 HF per arm, we would have 80% power in each
45 country to detect as significant a difference between a proportion of due vaccines given from 75% in
46 the control to 85% in the intervention arm, assuming one child per household, 30 households per HF
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 and a between-HF variation equivalent to a k of 0.25, where k is equal to the standard deviation
4
5 divided by the mean. The value of k is unknown, but was chosen in line with general observations by
6
7 Hayes and Bennet [20].
8

9
10 For data quality outcomes, with 35 HF per arm we would be able to detect as significant a difference
11
12 from a ratio of 0.7 (reported : recorded) vaccinations in the control arm to 0.8 with the intervention
13
14 with 80% power, assuming 100 recorded vaccinations per HF and a standard deviation of 0.25 in the
15
16 ratios between HF.
17

18
19 In terms of diagnostic scope, we would be able to detect an increase in the proportion of child-visits
20
21 with more than one diagnosis from 30% to 35% with 80% power with 35 HF per arm, 60 records per
22
23 HF and assuming a k of 0.25.
24

25
26 We would be able to detect as significant an increase from 50% of treatments having a
27
28 corresponding appropriate diagnosis to 60% with 80% power assuming 35 HF per arm, 1 treatment
29
30 per child, 25 children per HF and variation between HF corresponding to $k = 0.25$.
31

32
33 For the outcome related to health workers' satisfaction, we would be able to detect as significant an
34
35 increase from 50% of health workers satisfied to 75%, with 80% power assuming 35 HF, three health
36
37 workers per HF and a variation between HF equivalent to $k = 0.25$. Since this variable is measured in
38
39 the end-line survey only, we used the formula in Hayes and Bennet [20].
40

41
42 In summary, in each country we require 35 HF per arm, three HW per HF, 100 vaccination records
43
44 per HF, 60 sick child records per health facility and 30 children per health facility catchment area.
45
46
47

48 **Data collection and management**

49
50 Data collection took place at baseline and will take place again at the end of the study. Data is
51
52 collected from health facilities, from the households in the catchment areas of the included health
53
54 facilities and also from district offices.
55

56
57 For data quality and data use outcomes, HF registers, tallies and reports will be scrutinised. For
58
59 population based outcomes, we carry out household surveys at baseline and at end-line. We use
60

1
2
3 standard approaches for these types of surveys [21]. Households are visited, the research project is
4
5 briefly introduced and consent requested. Ideally, mothers of alive children or women in child-
6
7 bearing age were interviewed in order to obtain information on living children (i.e. vaccination
8
9 history) and death events, respectively, using home-based records if available and accessible.
10
11 Patients' satisfaction will be assessed using the PSQ-18 satisfaction questionnaire [22,23,24].
12
13 Essentially, the tool enables practitioners to investigate the extent to which their health care service
14
15 meets the perceived needs of their client group and pinpoint areas for improvement [24]. The
16
17 interview will be conducted with consenting patients as close to their care encounter as possible
18
19 [25]. Data tools are translated into the official languages of the study countries and pilot tested for
20
21 consistent meaning and relevance to the setting. Data collectors are also able to communicate in
22
23 local languages. The Satisfaction of Employees in Health Care (SEHC) survey is a validated tool to
24
25 assess staff satisfaction. It was first developed and validated in a low-income country (Ethiopia) [26]
26
27 and later successfully validated in a high-income country (USA) [27].
28
29

30
31 We use a mix of paper and electronic data (ODK [28]) collection tools. Data collectors are trained to
32
33 minimise error. Tools are piloted before implementing. ODK data is regularly stored and sent to
34
35 secure servers, as soon as data collectors reach their office base. Data from paper tools is double
36
37 entered and compared and sent to secure servers. Each data collection tool has its corresponding
38
39 electronic database that is cleaned and submitted to the analyses. All data is anonymised at the point
40
41 of data collection or as soon as possible in the data management process. Data is labelled with an
42
43 arm code (e.g. 'A' or 'B') without any further information allowing to disclose which data items
44
45 belong to the intervention or to the control arms, ensuring blinding during data analyses.
46
47

48
49 Quality will be assured through several mechanisms: piloting of data collection tools; thorough
50
51 training of field workers; checking missing data related; double, independent data entry from papers
52
53 into digital databases; early descriptive analyses to detect potential outliers; fieldworkers tracking
54
55 and supervision.
56
57
58
59
60

Data analysis

The analysis will be carried out for each country separately, and based on intention-to-treat.

Baseline population and health facility characteristics (i.e. basic demographic characteristics of population and health workers, professional profile of health workers, health facility size and services) will be summarised. If large imbalances are observed at baseline, the variables can be used to adjust the effect estimate comparisons [29,30].

The analyses vary for the different primary outcomes due to the unit of measurement and levels of clustering, the type of variable, and whether measurements were taken at baseline and endpoint or endpoint only. We use regression models to allow us to estimate the effect of the outcome while flexibly accounting for these issues and allowing adjustment for potential confounders.

Logistic regression will be used for the binary variables: vaccine adherence is measured by determining whether each vaccine due was received, and treatment appropriateness by whether each treatment was correctly prescribed. Data concordance and diagnostic scope are count variables and may be analysed with Poisson regression, depending on their distribution. The regression model for HW satisfaction will depend on how it is distributed.

The outcomes have different levels of clustering (children or consultations, HW, HF). We will account for these levels of clustering by including random effects in the regression models.

Four of the primary outcomes are measured at baseline and end-line. The effect of the intervention will be estimated using an interaction term between arm and survey in the regression models: i.e. is the change in the outcome between baseline and follow-up in the intervention arm different to the change between baseline and follow-up in the control arm? The effect of HW satisfaction, measured only at end-line, will be estimated as the difference between the intervention and control arm.

All estimates for the effect of the intervention will be presented with 95% confidence intervals. The analyses will be carried out using R [31].

Measures to minimise bias

Statistical analyses will be carried out blindly, without knowledge of what health facilities or population in the catchment area belong to the intervention or control groups. Only when the analysis code is considered as definitive and fixed, will results be shared with the wider investigators team and the arms for health facilities and population will be disclosed.

Outcome measurement bias may take place where data from the HIS, which is the focus of the intervention, is used to measure outcomes. However, we will minimise this by assessing population based outcomes at household level.

Contamination (i.e. the intervention affects individuals or units assigned to the control arm) may happen via the exchanges between health workers from health facilities in both arms; for example: in monthly district data quality meetings, managerial meetings; or through inputs from supervisors who influence control health facilities with intervention tips encountered in health facilities of the intervention arms. One mechanism to address this issue is using a district-based cluster randomisation scheme. However, we consider that (i) contamination, despite increasing the awareness of health works in control health facilities, will hardly influence the decision making mechanisms that the HIS intervention focuses on; and (ii) randomisation at the level of district poses additional challenges that are not worth the marginal benefit of reducing a doubtful contamination [32].

The spill-over effect (i.e. benefits of the intervention extend beyond their direct recipients) [33] may take place in higher levels of the health systems; e.g. district data managers and programme managers may experience the benefits of better structured and more timely data produced in health facilities in the intervention arms. The trial will have no capacity to quantitatively account for spill overs at higher levels of the system, due to the limited number of higher level administrative areas that will be involved in the trial. However, through process indicators, we will consider potential benefits and harms of the intervention at higher levels of the system.

1
2
3 A challenge is the Hawthorne effect (i.e. observer effect). Both health workers in the intervention
4 and in the control sites will have an awareness of being observed as data collection activities will be
5 at the same level of intensity in the two arms. Therefore, there should be no differential effect.
6
7

8
9
10 Analyses will be based on the intention-to-treat. It is important to closely monitor if the intervention
11 HFs consistently use the new HIS tools and approaches. The data collection team and the trial
12 monitoring team will check if old forms are still being used in the intervention health facilities.
13
14

15
16 However, we do not expect health facilities to migrate between intervention and control arms, or
17 vice versa, due to feasibility issues. On the other hand, some household members in a given
18 catchment area may decide to seek for health care in a health facility belonging to another trial arm.
19
20
21

22 In these cases, households will be analysed as belonging to the original trial arm.
23
24
25

26 **Ethics and dissemination**

27
28
29 Ethics committees in Côte d'Ivoire, Mozambique and Nigeria approved the study in their respective
30 countries. To date, some modifications to the protocol have taken place. In Côte d'Ivoire, we decided
31 to select study areas close to the research institution base on logistics and practical reasons, instead
32 of selecting an area in the north of the country, where poorer health indicators have been described.
33
34
35

36
37 In Mozambique, the low density of HF per population implied extremely vast distances between HF
38 and this, coupled with the rainy season, made the trial unfeasible in the originally selected Nampula
39 province. After consultations, we decided to move the trial to the province of Inhambane and cancel
40 the household survey. The allocation of HF to the intervention and control arms was completed using
41 random number generation.
42
43
44
45

46
47 We plan to disseminate the findings of the trials as one of the few examples of studies assessing the
48 effects of health information systems interventions using experimental study designs [34]. Most of
49 the experimental studies on HIS are circumscribed to specific health care areas (e.g. tuberculosis,
50 vaccination, cardiovascular disease) and very few have a system-wide approach (e.g. PHC) [34].
51
52
53
54
55

56
57 Experimental studies for health systems interventions are sometimes dismissed because of their
58
59
60

1
2
3 limited capacity to provide reliable explanations of complex health system issues [35]. While we
4
5 acknowledge these limitations, there is also a need for more robust evidence on the effects of these
6
7 types of health system interventions [36] and there are also good examples of experimental studies
8
9 reporting findings that can make it to the policy arena [37]. When embarking on this research, we
10
11 considered from the outset the type of evidence required to be disseminated and included into
12
13 systematic reviews [38], guidance development [39] and eventually recommendations for policy and
14
15 practice [40].
16
17
18
19
20

21 We acknowledge the challenges of carrying out research on health care provided to remote, rural
22
23 communities (in this case in Sub-Saharan Africa). However, it is only in these remote areas where
24
25 research about their specific problems and needs can take place. Challenges included long distances,
26
27 poor conditions of roads, unreliable communications and limited food and accommodation services,
28
29 all of them to be proactively handled to keep the quality of work and the morale of researchers and
30
31 collaborators. We expect that the dissemination of findings in meetings, conferences and
32
33 publications will contribute to a better understanding of what it takes to make research in
34
35 challenging contexts.
36
37
38

39 The engagement and ownership of partners within this research has also been instrumental in order
40
41 to plan and implement the CRCT. The intervention actually targets a governmental sub-system (the
42
43 HIS) for which we required more than permission but also endorsement and active support. We
44
45 achieved this level of collaboration by ensuring the participation of key stakeholders in key phases of
46
47 the whole project, from inception till the implementation of the last phases, through frequent
48
49 communication and workshops. The PHISICC programme includes targeted activities to keep
50
51 decision-makers engaged and we are planning to share the findings through workshops as well as
52
53 online and face-to-face events to disseminate the lessons learned from the trial and the whole
54
55 research.
56
57
58
59
60

1
2
3 We also expect that the dissemination of our findings among partners and competitors will
4
5 contribute to the current debates on the digitalisation of health information systems. WHO
6
7 recommendations on the matter are weak since the underlying evidence to support these
8
9 recommendations is inconsistent [41]. The principles and methodological approaches in PHISICC can
10
11 be applied to the development of any technological solution, being on paper, digital or mixed.
12
13
14 Finally, we expect that the results of the trials, both quantitative and qualitative, will be able to
15
16 inform policies on how to make HIS responsive to providers' decision-making needs, particularly in
17
18 health services where the most vulnerable live.
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Authors' contributions

XBC, AOI, AM, RBY and CAU prepared the proposal for the funding agency, conceived the study and produced the data collection tools. SG ensured the regulatory, ethical and trial monitoring components. AR developed the analytical approaches and made the sample size calculations. RBY, MS and SB adapted the protocol to the context of Côte d'Ivoire and managed the administrative and ethical approvals in the country; AOI, NE, ON, ANN and ABG likewise in Nigeria; AM, SML and GM, in Nampula province (Mozambique); JS and TM adapted the protocol and acquired ethical and administrative clearances for Inhambane province (Mozambique). DB is chair of the PHISICC Technical Advisory Group (TAG) and has coordinated multiple formal and informal inputs. LKK and DB have advised on the adequacy of the study protocol within the overall PHISICC proposal and TAG advices. All country teams participated in PHISICC workshops and ensured that the protocol was suitable to countries realities; developed data collection tools and training materials. They are responsible for the implementation of the trial in each country. XBC drafted the first version of the manuscript. All authors commented on several versions of the manuscript.

Acknowledgements

Members of the PHISICC Technical Advisory Group, chaired by David Brown: Blanche Anya (WHO AFRO), Abdallah Bchir (former Gavi), Marta Gacic-Dobo (WHO HQ), Richard Greffé (AIGA), Pamela Mitula (WHO AFRO), Sandy Oliver (UCL Institute) and Chris Wolf (BMGF).

Research collaborators: Momade Ali, Celso Belo, Bassirou Bonfoh, Lisa Diallo, John Ferreira, Bernard Guessanbi, Caitlin Jarrett, Inza Koné, Felix Malé, Kouadio M'bra, David O'Donnell, Damaris Rodríguez (Sonder Collective), Melanie Wendland and Meike Zuske (Swiss TPH).

Stakeholders from Côte d'Ivoire, Mozambique and Nigeria, participating in the consultation processes.

Data sharing statement

We report on a protocol of three Cluster Randomised Controlled Trials and therefore no data is available yet. On completion of the trials, access to data will be available from the national research institutions in Côte d'Ivoire, Mozambique and Nigeria, in publications and in the funding agency.

Individual participants' data on vaccination and antenatal care outcomes as well as health workers and users perceptions will be anonymised and made available via an online data repository for any purpose and access will be granted following a review of requests by the Swiss TPH contract officer. Data, with DOIs, will be made available during the second semester of 2021.

Available documents include study protocols, analytical plan, informed consent forms and analytical code.

Ethical approvals in Côte d'Ivoire, Mozambique and Nigeria

- Comité National Ethique des Sciences de la Vie et de la Santé (CNESVS), reference: 024-19/MSHP/CNESVS-kp (Côte d'Ivoire)
- Comité Institucional de Bioética para Saúde da Universidade Lúrio, reference: 16.2/Julho/CBISUL/19 (Mozambique)
- Secretary, Government of Cross River State of Nigeria, Ministry of Health, Calabar Health Research Ethics Committee, reference: CRS/MH/HREC/018/Vol. V1/151 (Nigeria)
- Ethikkommission Nordwest- und Zentralschweiz (EKNZ), reference: 2018-01059 (Switzerland).

Funding

Funded by the Bill & Melinda Gates Foundation, grant number INV-010193 / OPP1135947.

Abbreviations

BMGF Bill & Melinda Gates Foundation

CRCT Cluster Randomised Controlled Trial

sd Standard deviation

HF Health Facility

HW Health worker

WHO World Health Organisation

Additional files

- Additional file 1: CONSORT statement checklist.

Competing interests

No, there are no competing interests for any author.

References

- 1 WHO. Monitoring results with health facility information systems: A technical consultation Glion-sur-Montreux, Switzerland, 11–12 June 2014. WHO 2014.
- 2 WHO. Health facility & community data toolkit. WHO 2014.
- 3 Bosch-Capblanch X, Ronveaux O, Doyle V, Remedios V, Bchir A. Accuracy and quality of immunization information systems in forty-one low income countries. *Trop Med Int Health*. 2009 Jan;14(1):2-10. doi: 10.1111/j.1365-3156.2008.02181.x. PMID: 19152556.
- 4 Loder E. Promises, promises. *BMJ*. 2017 Mar 23;356:j1446. doi: 10.1136/bmj.j1446. PMID: 28336513.
- 5 Wenger N, Méan M, Castioni J, Marques-Vidal P, Waeber G, Garnier A. Allocation of Internal Medicine Resident Time in a Swiss Hospital: A Time and Motion Study of Day and Evening Shifts. *Ann Intern Med*. 2017 Apr 18;166(8):579-586. doi: 10.7326/M16-2238. Epub 2017 Jan 31. PMID: 28135724.
- 6 Bosch-Capblanch X, Zuske M, Auer C, Amadasoun S. Effects of interventions to improve ‘health information use systems’ on the quality and use of information. National Institute for Health Research. PROSPERO - https://www.crd.york.ac.uk/prospero/display_record.php?RecordID=35218.
- 7 Bosch-Capblanch X, Zuske M, Auer C. Making information support systems relevant to decision-makers: an effectiveness systematic review and a framework synthesis. Abstracts of the 25th Cochrane Colloquium, Edinburgh, UK. *Cochrane Database of Systematic Reviews* 2018;(9 Suppl 1). <https://doi.org/10.1002/14651858.CD201801>.
- 8 Schulz KF, Altman DG, Moher D, for the CONSORT Group. CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials.
- 9 Campbell MK, Piaggio G, Elbourne DR, Altman DG; for the CONSORT Group. Consort 2010 statement: extension to cluster randomised trials. *BMJ*. 2012;345:e5661.
- 10 R Core Team. R: A Language and Environment for Statistical Computing [Internet]. Vienna, Austria; 2020. Available from: <https://www.R-project.org/>
- 11 Bostoen K, Chalabi Z. Optimization of household survey sampling without sample frames. *Int J Epidemiol*. 2006 Jun;35(3):751-5. doi: 10.1093/ije/dyl019. Epub 2006 Feb 15. PMID: 16481364.
- 12 Johnson, Sandhu, Tyler, “The Next Step for Human-Centered Design in Global Public Health” *Stanford Social Innovation Review*, Oct. 22, 2019.
- 13 WHO. Global Vaccine Action Plan 2011-2020. WHO 2013.

- 1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
- 14 Carroli G, Villar J, Piaggio G, Khan-Neelofur D, Gülmezoglu M, Mugford M, Lumbiganon P, Farnot U, Bergsjø P; WHO Antenatal Care Trial Research Group. WHO systematic review of randomised controlled trials of routine antenatal care. *Lancet*. 2001 May 19;357(9268):1565-70. doi: 10.1016/S0140-6736(00)04723-1. PMID: 11377643.
- 15 Alpern R, Canavan ME, Thompson JT, McNatt Z, Tatek D, Lindfield T, Bradley EH. Development of a brief instrument for assessing healthcare employee satisfaction in a low-income setting. *PLoS One*. 2013 Nov 5;8(11):e79053. doi: 10.1371/journal.pone.0079053. PMID: 24223878; PMCID: PMC3818514.
- 16 Doll WJ and Torkzadeh G. The measurement of end-user computing satisfaction. *MIS Quarterly* 1988;12(2):259–74; 19.
- 17 Ilias A, Yasoa MR, Razaka MZA and Rahmana RA. The study of end-user computing satisfaction (EUCS) on computerised accounting system (CAS) among Labuan F.T. Government sectors: a case study in the responsibility centres. *Labuan e-Journal of Muamalat and Society* 2007;1:1–13.
- 18 Arifeen SE, Hoque DM, Akter T, Rahman M, Hoque ME, Begum K, Chowdhury EK, Khan R, Blum LS, Ahmed S, Hossain MA, Siddik A, Begum N, Sadeq-ur Rahman Q, Haque TM, Billah SM, Islam M, Rumi RA, Law E, Al-Helal ZA, Baqui AH, Schellenberg J, Adam T, Moulton LH, Habicht JP, Scherpbier RW, Victora CG, Bryce J, Black RE. Effect of the Integrated Management of Childhood Illness strategy on childhood mortality and nutrition in a rural area in Bangladesh: a cluster randomised trial. *Lancet*. 2009 Aug 1;374(9687):393-403. doi: 10.1016/S0140-6736(09)60828-X. PubMed PMID: 19647607.
- 19 Tilahun H, Fekadu B, Abdisa H, Canavan M, Linnander E, Bradley EH, Berman P. Ethiopia's health extension workers use of work time on duty: time and motion study. *Health Policy Plan*. 2017 Apr 1;32(3):320-328. doi: 10.1093/heapol/czw129. PubMed PMID: 27658649.
- 20 Hayes RJ, Bennett S. Simple sample size calculation for cluster-randomized trials. *Int J Epidemiol*. 1999 Apr;28(2):319-26.
- 21 The DHS program. Demographic and Health surveys. <https://dhsprogram.com/>.
- 22 Thayaparan, A. a. (2013). The Patient Satisfaction Questionnaire Short Form (PSQ-18) as an adaptable, reliable, and validated tool for use in various settings. *Med Educ Online*, 18.
- 23 The Health Foundation. (2013, June). Measuring patient experience (no.18). Retrieved Nov 6, 2018, from www.health.org.uk: <https://www.health.org.uk/publication/measuring-patient-experience>.

- 1
2
3
4
5
6
7 24 Grogan, S. C. (2000). Validation of a questionnaire measuring patient satisfaction with general
8 practitioner services. *Quality in Health Care*, 9, 210-215.
- 9
10 25 LaVela, S. L. (2014). Evaluation and measurement of patient experience. *Patient Experience*
11 *Journal*, 1(1), Article 5.
- 12
13 26 Alpern, R. C. (2013, November). Development of a Brief Instrument for Assessing Healthcare
14 Employee Satisfaction in a Low-Income Setting. *PLOS ONE*, 8(11), e79053.
- 15
16 27 Chang, E. C. (2017). Measuring job satisfaction among healthcare staff in the United States: a
17 confirmatory factor analysis of the Satisfaction of Employees in health Care (SEHC) survey.
18 *International Journal for Quality in health Care*, 29(2), 262-268.
- 19
20 28 Open Data Kit. <https://opendatakit.org/>. Accessed 5th February 2021.
- 21
22 29 Pocock S, et al. Subgroup analysis, covariate adjustment and baseline comparisons in clinical trial
23 reporting: current practice and problems. *Stat Med*. 2002;21:2917-30.
- 24
25 30 The CONSORT Group. CONSORT Statement [Internet]; item 15.
- 26
27 31 R Core Team. R: A language and environment for statistical computing. R Foundation for
28 Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.
- 29
30 32 Torgerson David J. Contamination in trials: is cluster randomisation the answer? *BMJ* 2001;
31 322:355.
- 32
33 33 Benjamin-Chung J, Arnold BF, Berger D, Luby SP, Miguel E, Colford JM Jr, Hubbard AE. Spillover
34 effects in epidemiology: parameters, study designs and methodological considerations. *Int J*
35 *Epidemiol*. 2018 Feb 1;47(1):332-347. doi: 10.1093/ije/dyx201. PubMed PMID: 29106568; PubMed
36 Central PMCID: PMC5837695.
- 37
38 34 Leon N, Balakrishna Y, Hohlfeld A, Odendaal WA, Schmidt BM, Zweigenthal V, Anstey Watkins J,
39 Daniels K. Routine Health Information System (RHIS) improvements for strengthened health system
40 management. *Cochrane Database Syst Rev*. 2020 Aug 13;8:CD012012. doi:
41 10.1002/14651858.CD012012.pub2. PMID: 32803893.
- 42
43 35 Allanson ER, Tunçalp Ö, Vogel JP, Khan DN, Oladapo OT, Long Q, Gülmezoglu AM. Implementation
44 of effective practices in health facilities: a systematic review of cluster randomised trials. *BMJ Glob*
45 *Health*. 2017 Jul 20;2(2):e000266. doi: 10.1136/bmjgh-2016-000266. PMID: 29081997; PMCID:
46 PMC5656132.
- 47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3
4
5
6
7 36 Brown DW, Bosch-Capblanch X, Shimp L. Where Do We Go From Here? Defining an Agenda for
8 Home-Based Records Research and Action Considering the 2018 WHO Guidelines. *Glob Health Sci*
9 *Pract.* 2019 Mar 29;7(1):6-11. doi: 10.9745/GHSP-D-18-00431. PMID: 30877139; PMCID:
10 PM6538131.

11
12
13
14 37 Joos O, Silva R, Amouzou A, Moulton LH, Perin J, Bryce J, et al. Evaluation of a mhealth data
15 quality intervention to improve documentation of pregnancy outcomes by health surveillance
16 assistants in Malawi: a cluster randomized trial. *PloS One* 2016;11(1):e0145238. PMID: 26731401.
17 DOI: <https://doi.org/10.1371/journal.pone.0145238>

18
19
20
21 38 Langlois EV, Ranson MK, Bärnighausen T, Bosch-Capblanch X, Daniels K, El-Jardali F, Ghaffar A,
22 Grimshaw J, Haines A, Lavis JN, Lewin S, Meng Q, Oliver S, Pantoja T, Straus S, Shemilt I, Tovey D,
23 Tugwell P, Waddington H, Wilson M, Yuan B, Røttingen JA. Advancing the field of health systems
24 research synthesis. *Syst Rev.* 2015 Jul 10;4:90. doi: 10.1186/s13643-015-0080-9. PMID: 26159806;
25 PMCID: PMC4498528.

26
27
28
29 39 Bosch-Capblanch X, Lavis JN, Lewin S, Atun R, Røttingen JA, Dröschel D, Beck L, Abalos E, El-Jardali
30 F, Gilson L, Oliver S, Wyss K, Tugwell P, Kulier R, Pang T, Haines A. Guidance for evidence-informed
31 policies about health systems: rationale for and challenges of guidance development. *PLoS Med.*
32 2012;9(3):e1001185. doi: 10.1371/journal.pmed.1001185. Epub 2012 Mar 6. PMID: 22412356;
33 PMCID: PMC3295823.

34
35
36
37 40 WHO. WHO recommendations on home-based records for maternal, newborn and child health.
38 WHO 2018.

39
40
41
42 41 WHO. Recommendations on digital interventions for health system strengthening. WHO
43 guideline. WHO 2019.
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Tables and Figures

Table 1. Comparison of new (intervention) and current (control) tools.

Characteristics	New (PHISICC intervention) tools	Old (control tools)
Development approach	Human Centred Design, co-creation with users	Centrally done, based on data and information experts
Visual language	Standardised across tools	No visual elements
Information structure	Following clinical processes	Tabular form, following reporting requirements
Decision aids	Icons representing mild, moderate and severe conditions	Not available
Register books layout	Landscape, DIN-A3	Depending on health care area; often much larger than DIN-A3
Tally sheets to aid counting events	For each health care area, to be filled as health care events take place	Only for vaccination, to be filled as vaccinations take place
Reporting	Integrated with tallying / counting	Requires revisiting register books at the end of the month

Table 2. Outcomes and parameters used to estimate the sample sizes.

	Outcome name	Subjects	Definition	Baseline estimate	Expected change	Comments
1	Vaccination adherence	Children under-1 in (sample of households in catchment areas)	Number of vaccines given in the previous calendar year over the number of vaccine due in the same period	75 given per 100 due	Increase of 10 per 100	Vaccines are clustered within children, and children within HFs
2	Data concordance	Recording tools in health facilities (samples of records)	Number of health care events (e.g. vaccinations, antenatal care consultations) recounted in the previous calendar year versus the number of health care events reported in the same time period	7 recounted for each 10 reported [3]	Increase of 1 recounted	A single estimate can be obtained in each HF or by time periods (no clustering)
3	Diagnostic scope	Records of sick child consultations (samples of records)	Number of diagnosis in each sick child consultation during the previous calendar year	30% with more than 1 diagnosis	35% with more than 1 diagnosis	Individual consultations are clustered within HF
4	Treatment appropriateness	Records of sick child consultations (samples of records)	Number of treatments correctly prescribed in each sick child consultation during the previous calendar year	Half appropriate over all consultations	Increase to 60%	Individual consultations are clustered within HF (one treatment per child)
5	Health workforce satisfaction	Health workers (all health workers form include health facilities)	Degree (score) of satisfaction across all health facilities in each arm, with the intervention	50% satisfied	75% satisfied	Maybe two or three health workers can be approached in each health facility

Figure 1. PHISICC research programme structure, processes, deliverables and flow of evidence.

Footnote to Figure 1. WS: work stream. Timelines are approximate.

1
2
3 **Figure 2. CONSORT diagram: trial flow chart.**
4
5

6 Separate file.
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

For peer review only

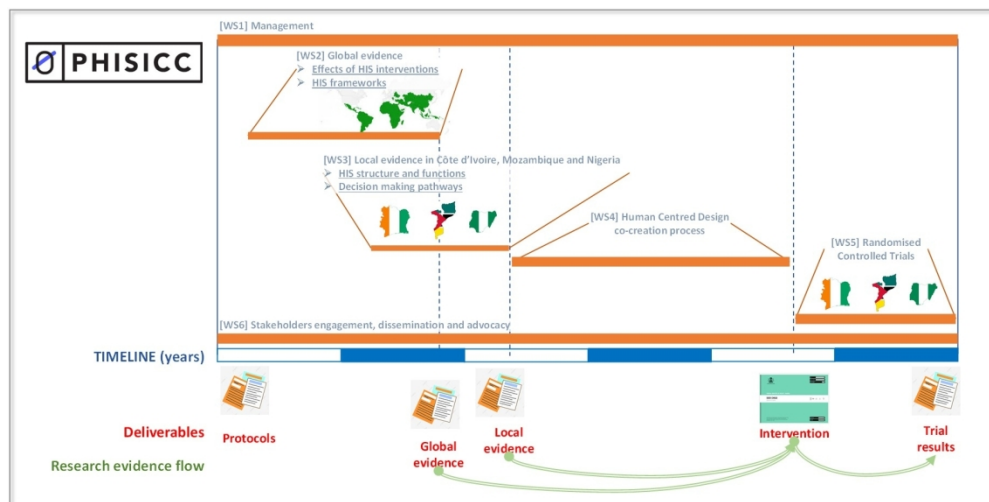


Figure 1. PHISICC research programme structure, processes, deliverables and flow of evidence.

312x157mm (150 x 150 DPI)

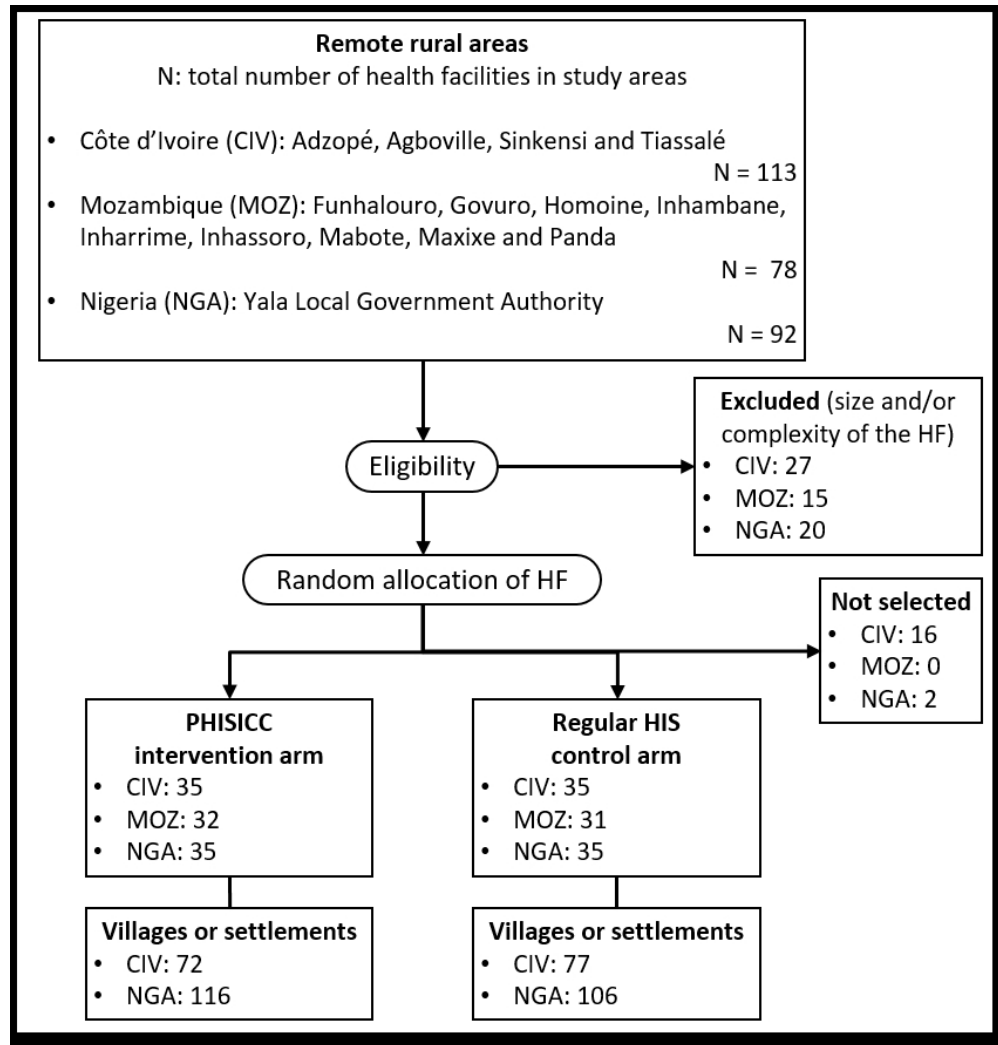


Figure 2. CONSORT diagram: trial flow chart.

144x150mm (150 x 150 DPI)

```
1
2
3 #
4 # clustersampleSize_proportions_baseline&endline.r
5 # get power of cluster randomised trial for binary outcomes (baseline and
6 # endline surveys)
7 # 2 groups (control & intervention)
8 # clustered within HF
9
10
11 rm(list=ls())
12
13 # if the package lme4 is not already installed (needed for regression with
14 # random effects)
15 # install.packages(lme4)
16 require(lme4)
17 #install.packages("reshape")
18 library(reshape)
19
20
21
22 # INPUTS
23 numGroups<-2
24 numHFPerGroup<-35
25 numTrialsToSimulate<-100
26 # numTrialsToSimulate: use 10 to test that the script runs, use 100 or 1000 for
27 # precise estimate of power
28
29
30
31
32 # choose input set and remove #s to run
33
34 # inputs for 'treatments with appropriate diagnosis'
35 pInterv<-0.60
36 pControl<-0.50
37 sdHFcluster<-0.55
38 # for k=0.1, 0.20; for k=0.25, 0.55
39 numObsPerHF<-25
40
41 # inputs for vaccination adherence
42 # proportions in interventions and control groups
43 # pInterv<-0.8
44 # pControl<-0.75
45 # sdHFcluster<-2.63
46 # numObsPerHF<-30
47
48
49
50 # inputs for 'more than one diagnosis'
51 # pInterv<-0.35
52 # pControl<-0.30
53 # sdHFcluster<-0.39
54 # for k=0.1, 0.16; for k=0.25, 0.39
55 # numObsPerHF<-60
56
57
58
59 # NB getsd is a function at the bottom of the script to turn k into sdHFcluster
60 (sdHFcluster is on the logit scale)
```

```

1
2
3
4
5
6 # --- simulation ----
7
8 # SET UP DATA STRUCTURE (intervention, HF)
9 totNumHF <- numHFPerGroup*numGroups
10 HFList<-seq(1:totNumHF)
11 interv<- rep(c(0,1),each=(totNumHF/2) )
12 intervEffect<-rep( c(0,(log(pInterv/(1-pInterv)) -
13 log(pControl/(1-pControl))) ), each=(totNumHF/2) )
14
15
16 xtemp<-cbind(interv,HFList,intervEffect)
17
18 # SET UP STORE FOR PVALUES AND PRECISION
19 storeResults<-array(-9,dim=c(numTrialsToSimulate,3))
20 colnames(storeResults)<-c("pvalue","coeff","stderr")
21
22
23 # LOOP THROUGH THE SIMULATIONS
24
25 for (i in 1:numTrialsToSimulate) {
26
27 # simulate the HF cluster effects
28 HFEffect<-rnorm(totNumHF,mean=0,sd=sdHFcluster)
29 xtemp2a<-cbind(xtemp, HFEffect)
30 xtemp2a<-data.frame(xtemp2a)
31
32
33 # get expected proportions (pre and post)
34 xtemp2a$expectedprelogodds<-log(pControl/(1-pControl)) + xtemp2a$HFEffect
35
36 xtemp2a$expectedpostlogodds<-log(pControl/(1-pControl)) +
37 xtemp2a$intervEffect + xtemp2a$HFEffect
38
39 xtemp2a$expectedpre<-exp(xtemp2a$expectedprelogodds)/(1+exp(xtemp2a$expectedpre
40 logodds))
41
42 xtemp2a$expectedpost<-exp(xtemp2a$expectedpostlogodds)/(1+exp(xtemp2a$expectedp
43 ostlogodds))
44
45 # expand by the number of observations per HF
46 xtemp2b<-untable(xtemp2a, num=numObsPerHF)
47 numObs<-dim(xtemp2b)[1]
48
49
50 # simulate individual observations from cluster mean rates
51 simObsPost<-rep(0,numObs)
52 simObsPre<-rep(0,numObs)
53 for (j in 1:numObs) {
54 simObsPost[j]<-rbinom(n=1, size=1,prob=xtemp2b$expectedpost[j])
55 simObsPre[j]<-rbinom(n=1, size=1,prob=xtemp2b$expectedpre[j])
56 }
57 # drop variables not needed further
58 xtemp2b$expectedpostlogodds<-NULL; xtemp2b$expectedprelogodds<-NULL
59
60

```



```
1
2
3
4 # stack pre and post observations
5 # get post
6 xtemp3<-cbind(xtemp2b,simObsPost)
7 xtemp3<-data.frame(xtemp3)
8 xtemp3$simObs<-xtemp3$simObsPost
9 xtemp3$simObsPost<-NULL
10 xtemp3$post<-1
11 # get pre
12 xtemp4<-cbind(xtemp2b,simObsPre)
13 xtemp4<-data.frame(xtemp4)
14 xtemp4$simObs<-xtemp4$simObsPre
15 xtemp4$simObsPre<-NULL
16 xtemp4$post<-0
17 xtemp4$interv<-0
18 xtemp5<-rbind(xtemp3,xtemp4)
19
20
21
22 # carry out analysis for individual trial
23 m <- glmer(simObs ~ as.factor(interv) + post + (1 | HFList),
24 data<-xtemp5, family=binomial)
25
26 # store result of individual trial in storeResults (p-value, coefficient
27 and std error)
28 out1<-summary(m)$coefficients
29 storeResults[i,2]<-out1[2,1]
30 storeResults[i,3]<-out1[2,2]
31 storeResults[i,1]<-out1[2,4]
32
33
34 print(i)
35
36 } # End of loop
37
38 # calculate power
39 pvalue<-storeResults[,1]
40 power<-length(pvalue[pvalue<0.05])/length(pvalue)
41
42 cat("power ", power, "\n")
43
44
45
46
47 # ----- run to here -----
48
49
50
51
52
53
54
55 # -----
56 # getsd: function to estimate between-cluster variation from k (Hayes and
57 Bennet sd/mean) and input base proportion (base0p)
58
59
60
```

```
1
2
3  getsd<-function(base0p,k){
4      sdcluster<-k*base0p
5      clusterEffect<-rnorm(1000,mean=0,sd=sdcluster)
6      expectedp<-base0p + clusterEffect
7      expectedp[expectedp>1]<-0.9999
8      expectedp[expectedp<0]<-0.0001
9      logitexpectedp<-log((expectedp)/(1-expectedp))
10     sdlog<-sd(logitexpectedp)
11     cat("estimated sdlog ", sdlog, "\n")
12 }
13
14
15 # example
16 getsd(0.30,0.25)
17
18 getsd(0.50, 0.25)
```

```
1
2
3 #
4 # clusterSampleSize_concordance.r
5 # get power of cluster randomised trial
6 # ratios (outcome is continuous)
7 # fixed to 2 groups
8 # records and reports clustered within HF
9 #
10
11
12 # if the package lme4 is not already installed (needed for regression with
13 random effects)
14 # install.packages(lme4)
15 # install.packages(lmerTest)
16 require(lme4)
17 require(lmerTest)
18
19
20
21 # EXAMPLE INPUTS
22 numGroups<-2
23 numHFPerGroup<-35
24 numReportedPerHF<-100
25 # assuming equal numbers of vaccinations per HF
26 numTrialsToSimulate<-100
27 # 100 or 1000 needed for precision of the power estimate, use 10 for test runs
28
29 ratioControl<-0.7
30 ratioInterv<-0.8
31 sdHFcluster<-0.25*0.8
32 # sdHFcluster is on the log scale, calculated using k=0.25
33
34
35
36
37 # --- run simulation from here ----
38
39 # SET UP DATA STRUCTURE (intervention, HF)
40 totNumHF<-numGroups*numHFPerGroup
41 HFList<-rep(seq(1:(numHFPerGroup*numGroups)),each=1)
42 interv<-c( rep(c(0,1),each=(totNumHF/2)))
43 intervEffect<-rep( c(0,(ratioInterv - ratioControl )), each=(totNumHF/2) )
44 xtemp<-cbind(interv,HFList,intervEffect)
45
46
47 # SET UP STORE FOR PVALUES AND PRECISION
48 storeResults<-array(-9,dim=c(numTrialsToSimulate,3))
49 colnames(storeResults)<-c("pvalue","coeff","stderr")
50
51
52 # LOOP THROUGH THE SIMULATIONS
53
54 for (i in 1:numTrialsToSimulate) {
55
56   # simulate the HF cluster effects
57
58   HFEffect<-rnorm(numHFPerGroup*numGroups,mean=0,sd=sdHFcluster)
59   xtemp2<-cbind(xtemp, HFEffect)
60
```

```

1
2
3
4     # get expected ratios (pre and post)
5     expectedpreratio<-ratioControl + HFEeffect
6     expectedpostratio<-ratioControl + intervEffect + HFEeffect
7     expectedpreratio[expectedpreratio<0.0001]<-0.0001
8     expectedpostratio[expectedpostratio<0.0001]<-0.0001
9
10    # simulate individual observations as poisson rate of number reported per
11    1 recorded
12    simObsPost<-rep(0,length(expectedpostratio))
13    simObsPre<-rep(0,length(expectedpreratio))
14    for (j in 1:length(expectedpostratio)) {
15        simObsPost[j]<-rpois(n=1,expectedpostratio[j]*numReportedPerHF)
16        simObsPre[j]<-rpois(n=1,expectedpreratio[j]*numReportedPerHF)
17    }
18
19
20
21    # stack pre and post observations
22    # post
23    xtemp3<-cbind(xtemp2,simObsPost)
24    xtemp3<-data.frame(xtemp3)
25    xtemp3$simObs<-xtemp3$simObsPost
26    xtemp3$simObsPost<-NULL
27    xtemp3$post<-1
28    # pre
29    xtemp4<-cbind(xtemp2,simObsPre)
30    xtemp4<-data.frame(xtemp4)
31    xtemp4$simObs<-xtemp4$simObsPre
32    xtemp4$simObsPre<-NULL
33    xtemp4$post<-0
34    xtemp4$interv<-0
35    # stack pre and post
36    xtemp5<-rbind(xtemp3,xtemp4)
37    xtemp5$distanceToOne<-abs(1-(xtemp5$simObs/numReportedPerHF))
38
39
40
41    # carry out analysis for individual trial
42    m <- lmer(distanceToOne ~ as.factor(interv) + post + (1|HFList),
43    data=xtemp5)
44
45    # store result of individual trial in storeResults (p-value, coefficient
46    and std error)
47    out1<-summary(m)$coefficients
48    # estimate
49    storeResults[i,2]<-out1[2,1]
50    # se
51    storeResults[i,3]<-out1[2,2]
52    # p-value
53    storeResults[i,1]<-out1[2,5]
54
55    print(i)
56
57
58 } # End of loop
59
60

```

```
1
2
3
4 # calculate power
5 pvalue<-storeResults[,1]
6 power<-length(pvalue[pvalue<0.05])/length(pvalue)
7
8 cat("power ", power, "\n")
9
10
11 # -----
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
```

Table 1: CONSORT 2010 checklist of information to include when reporting a cluster randomised trial

Section/Topic	Item No	Standard Checklist item	Extension for cluster designs	Page No *
Title and abstract				
	1a	Identification as a randomised trial in the title	Identification as a cluster randomised trial in the title	1
	1b	Structured summary of trial design, methods, results, and conclusions	See table 2	5
Introduction				
Background and objectives	2a	Scientific background and explanation of rationale	Rationale for using a cluster design	4
	2b	Specific objectives or hypotheses	Whether objectives pertain to the cluster level, the individual participant level or both	5
Methods				
Trial design	3a	Description of trial design (such as parallel, factorial) including allocation ratio	Definition of cluster and description of how the design features apply to the clusters	6
	3b	Important changes to methods after trial commencement (such as eligibility criteria), with reasons		Not applicable
Participants	4a	Eligibility criteria for participants	Eligibility criteria for clusters	6
	4b	Settings and locations where the data were collected		6
Interventions	5	The interventions for each group with sufficient details to allow replication, including how and when they were actually administered	Whether interventions pertain to the cluster level, the individual participant level or both	8
Outcomes	6a	Completely defined pre-specified primary and secondary outcome measures, including how and when they were assessed	Whether outcome measures pertain to the cluster level, the individual participant level or both	10 and Table 1
	6b	Any changes to trial outcomes after the trial commenced, with reasons		Not applicable
Sample size	7a	How sample size was determined	Method of calculation, number of clusters(s) (and whether equal or unequal cluster sizes are assumed), cluster size, a coefficient of intracluster correlation (ICC or k), and an indication of its uncertainty	12
	7b	When applicable, explanation of any interim analyses and stopping guidelines		Not yet applicable

Section/Topic	Item No	Standard Checklist item	Extension for cluster designs	Page No *
Randomisation:				
Sequence generation	8a	Method used to generate the random allocation sequence		7, 8
	8b	Type of randomisation; details of any restriction (such as blocking and block size)	Details of stratification or matching if used	Not applicable
Allocation concealment mechanism	9	Mechanism used to implement the random allocation sequence (such as sequentially numbered containers), describing any steps taken to conceal the sequence until interventions were assigned	Specification that allocation was based on clusters rather than individuals and whether allocation concealment (if any) was at the cluster level, the individual participant level or both	7
Implementation	10	Who generated the random allocation sequence, who enrolled participants, and who assigned participants to interventions	Replace by 10a, 10b and 10c	7
	10a		Who generated the random allocation sequence, who enrolled clusters, and who assigned clusters to interventions	7
	10b		Mechanism by which individual participants were included in clusters for the purposes of the trial (such as complete enumeration, random sampling)	10
	10c		From whom consent was sought (representatives of the cluster, or individual cluster members, or both), and whether consent was sought before or after randomisation	8
Blinding	11a	If done, who was blinded after assignment to interventions (for example, participants, care providers, those assessing outcomes) and how		8
	11b	If relevant, description of the similarity of interventions		Not applicable
Statistical methods	12a	Statistical methods used to compare groups for primary and secondary outcomes	How clustering was taken into account	15
	12b	Methods for additional analyses, such as subgroup analyses and adjusted analyses		15

Section/Topic	Item No	Standard Checklist item	Extension for cluster designs	Page No *
Results				Not yet applicable (protocol manuscript)
Discussion				
Limitations	20	Trial limitations, addressing sources of potential bias, imprecision, and, if relevant, multiplicity of analyses		18
Generalisability	21	Generalisability (external validity, applicability) of the trial findings	Generalisability to clusters and/or individual participants (as relevant)	18,19
Interpretation	22	Interpretation consistent with results, balancing benefits and harms, and considering other relevant evidence		Not yet applicable
Other information				
Registration	23	Registration number and name of trial registry		1
Protocol	24	Where the full trial protocol can be accessed, if available		1
Funding	25	Sources of funding and other support (such as supply of drugs), role of funders		20

* Page numbers: as seen in the document "draft_Proof_hi.pdf" (which has 34 pages)

CONSORT 2010 checklist of information to include when reporting a cluster randomised trial

Section/Topic	Item No	Standard Checklist item	Extension for cluster designs	Page No *
Title and abstract				
	1a	Identification as a randomised trial in the title	Identification as a cluster randomised trial in the title	1
	1b	Structured summary of trial design, methods, results, and conclusions	See table 2	5
Introduction				
Background and objectives	2a	Scientific background and explanation of rationale	Rationale for using a cluster design	4
	2b	Specific objectives or hypotheses	Whether objectives pertain to the cluster level, the individual participant level or both	5
Methods				
Trial design	3a	Description of trial design (such as parallel, factorial) including allocation ratio	Definition of cluster and description of how the design features apply to the clusters	6
	3b	Important changes to methods after trial commencement (such as eligibility criteria), with reasons		Not applicable
Participants	4a	Eligibility criteria for participants	Eligibility criteria for clusters	6
	4b	Settings and locations where the data were collected		6
Interventions	5	The interventions for each group with sufficient details to allow replication, including how and when they were actually administered	Whether interventions pertain to the cluster level, the individual participant level or both	8
Outcomes	6a	Completely defined pre-specified primary and secondary outcome measures, including how and when they were assessed	Whether outcome measures pertain to the cluster level, the individual participant level or both	10 and Table 1
	6b	Any changes to trial outcomes after the trial commenced, with reasons		Not applicable
Sample size	7a	How sample size was determined	Method of calculation, number of clusters(s) (and whether equal or unequal cluster sizes are assumed), cluster size, a coefficient of intracluster correlation (ICC or k), and an indication of its uncertainty	12
	7b	When applicable, explanation of any interim analyses and stopping guidelines		Not yet applicable

Section/Topic	Item No	Standard Checklist item	Extension for cluster designs	Page No *
Randomisation:				
Sequence generation	8a	Method used to generate the random allocation sequence		7, 8
	8b	Type of randomisation; details of any restriction (such as blocking and block size)	Details of stratification or matching if used	Not applicable
Allocation concealment mechanism	9	Mechanism used to implement the random allocation sequence (such as sequentially numbered containers), describing any steps taken to conceal the sequence until interventions were assigned	Specification that allocation was based on clusters rather than individuals and whether allocation concealment (if any) was at the cluster level, the individual participant level or both	7
Implementation	10	Who generated the random allocation sequence, who enrolled participants, and who assigned participants to interventions	Replace by 10a, 10b and 10c	7
	10a		Who generated the random allocation sequence, who enrolled clusters, and who assigned clusters to interventions	7
	10b		Mechanism by which individual participants were included in clusters for the purposes of the trial (such as complete enumeration, random sampling)	10
	10c		From whom consent was sought (representatives of the cluster, or individual cluster members, or both), and whether consent was sought before or after randomisation	8
Blinding	11a	If done, who was blinded after assignment to interventions (for example, participants, care providers, those assessing outcomes) and how		8
	11b	If relevant, description of the similarity of interventions		Not applicable
Statistical methods	12a	Statistical methods used to compare groups for primary and secondary outcomes	How clustering was taken into account	15
	12b	Methods for additional analyses, such as subgroup analyses and adjusted analyses		15

Section/Topic	Item No	Standard Checklist item	Extension for cluster designs	Page No *
Results				Not yet applicable (protocol manuscript)
Discussion				
Limitations	20	Trial limitations, addressing sources of potential bias, imprecision, and, if relevant, multiplicity of analyses		18
Generalisability	21	Generalisability (external validity, applicability) of the trial findings	Generalisability to clusters and/or individual participants (as relevant)	18,19
Interpretation	22	Interpretation consistent with results, balancing benefits and harms, and considering other relevant evidence		Not yet applicable
Other information				
Registration	23	Registration number and name of trial registry		1
Protocol	24	Where the full trial protocol can be accessed, if available		1
Funding	25	Sources of funding and other support (such as supply of drugs), role of funders		20