An educational intervention to enable children to assess claims about the benefits and harms of treatments

Allen Nsangi

PhD Summary

Faculty of Medicine

University of Oslo

1 TABLE OF CONTENTS

2 SUMMARY

Norsk sammendrag

Et skolebasert tiltak for å sette barn i stand til å vurdere påstander om positive og negative virkninger av helsetiltak

Bakgrunn

Vi møter påstander om effekt av tiltak overalt: Hva virker og hva virker ikke? Hva er bra for helsa og hva er skadelig? Ofte er påstandene ikke til å stole på, men mange klarer ikke å skille mellom det som er troverdig, og det som ikke er det. Å basere beslutninger på tynt funderte påstander medfører dårlige personlige valg, unødig lidelse, sløsing med tid og penger, og uhensiktsmessige politiske avgjørelser.

I Informed Health Choices-prosjektet utviklet vi undervisningsmateriell for å lære barn kritisk vurdering av helsepåstander. Formålet var å gjøre dem i stand til å ta informerte personlige valg, delta i den offentlig debatten og til å bidra til informerte helsepolitiske beslutninger når de blir voksne – som pasienter, helsepersonell, helsepolitikere og samfunnsborgere.

Forskningsmål

Det første målet var å utvikle undervisningsmateriell som skulle lære barneskoleelever i Uganda kritisk vurdering helsepåstander og å ta informerte helsebeslutninger. Det andre målet var å evaluere virkningen av materiellet på barnas evne til kritisk tenkning om helsepåstander og informerte beslutninger. Det tredje målet var å utforske mulige uforutsette konsekvenser og faktorer som kan innvirke på muligheten til å skalere opp bruken av materiellet. Det fjerde målet var å avklare om barna husket det de hadde lært i minst ett år.

Metoder

Vi utviklet læringsmateriellet gjennom en "human-centred design"-tilnærming. Vi utførte et randomisert forsøk med 120 skoler for å evaluere virkningene av materiellet på barnas evne til kritisk vurdering av helsepåstander og til å ta informerte beslutninger. Må målte resultater både like etter at de hadde brukt materiellet i undervisningen, og på nytt etter et år. Parallelt med forsøket gjorde vi en prosessevaluering, med strukturerte observasjoner i klasserommene, intervjuer og fokusgruppediskusjoner.

Funn

Under utviklingsprosessen registrerte vi at lærerne og elevene opplevde materiellet som nyttig, lett å bruke, og anvendelig i deres kontekst.

Det randomiserte forsøket viste at materiellet hadde en betydelig effekt på barnas evne til å vurdere helsepåstander, og at de beholdt læringsutbyttet i minst ett år.

Prosessevalueringen viste at de fleste deltakerne vurderte læringsmateriellet som viktig og relevant.

Konklusjon

Det er mulig å lære barn å tenke kritisk om helsepåstander og helsevalg, og å gjøre det i stor skala i et lavinntektsland.

Denne avhandlingen er en del av det internasjonale Informed Health Choices-prosjektet. (https://www.informedhealthchoices.org/)

English summary

An educational intervention to enable children to assess claims about the benefits and harms of treatments

Background

Claims about treatment effects are everywhere: What works, and what does not? What improves our health, what is harmful? Many such claims are unreliable and often times, people are unable to distinguish reliable from unreliable claims. Acting on unsubstantiated claims results in poor choices, unnecessary suffering, waste, and poorly informed policy decisions.

The Informed Health Choices Project aimed to address this problem by teaching children to think critically about health claims and choices. The goal was to enable them to make informed personal choices, participate in policy debates and contribute to informed health policy decisions when they are older – as patients, future health professionals, future policy makers and citizens.

Research Objectives

My first objective was to develop resources that teach primary school children in Uganda to assess health claims and make informed health choices. The second objective was to evaluate the impact of the resources on children's ability to think critically about health claims and choices. The third objective was to explore unintended consequences and factors that might affect scaling up use of the resources. The fourth objective was to ascertain if children retained what they learned for at least one year.

Methods

We employed a human-centred design approach to develop the resources. We carried out a cluster-randomized trial with 120 schools to evaluate the effect of the use of the resources on the children's ability to think critically about health claims and choices after using the resources and again after one year. We conducted a process evaluation alongside the trial, using structured classroom observations, interviews and focus group discussions.

Findings

Findings from the development process indicated that teachers and children found the resources useful, easy to use and appropriate for their context.

The randomized trial demonstrated a large effect on the children's ability to assess treatment claims and they retained what they learned for at least one year.

The process evaluation showed that most of the participants found the teaching resources important and relevant.

Conclusions

It is possible to teach children to think critically about health claims and choices on a large scale in a low- income country.

This thesis was part of the international Informed Health Choices Project (https://www.informedhealthchoices.org/).

3 DEDICATION

For my daughter, Elizabeth!

4 ACKNOWLEDGEMENTS

This thesis is based on work funded by the Norwegian Research Council, carried out as part of a collaboration between the Informed Health Choices Centre at the Norwegian Institute of Public Health in Oslo, Norway, Makerere University College of Health Sciences in Kampala, Uganda, Great Lakes University of Kisumu, Kenya, University of Rwanda, Kigali, Rwanda, Infodesign Lab, Oslo, Norway, and the James Lind Institute, Oxford, United Kingdom.

My greatest thanks to all the primary school teachers and children who made doing this work fun and worthwhile. Working with you, visiting your schools and attending some of your lessons felt like being back at primary school all over again and thank you for providing me that opportunity in adulthood.

In gratitude for their great help and support, I would also like to thank my main supervisor, Atle Fretheim, my co-supervisors, Andy Oxman and Nelson Sewankambo. Without your vision, encouragement and unwavering support, this thesis would not be accomplished. I consider myself fortunate for having had the opportunity to participate in many enlightening discussions with you during the course of this research. Standing on the shoulders of these giants has left me forever indebted to them.

In addition, I would like to thank several members of the Informed Health Choices research group; (Astrid Austvoll-Dahlgren, Sarah Rosenbaum, Matt Oxman, Angela Morelli, Claire Glenton, Simon Lewin, Margaret Kaseje, Laetitia Nyirazinyoye and Iain Chalmers), who have made tremendous contribution in various capacities to this great body of work, this thesis would not have been possible without your effort.

I would also like to thank my friend and colleague, Daniel Semakula for his daily support, through both the good and bad times, which meant a lot to me both personally and professionally. Special thanks to the entire Ugandan support team; Esther Nakyejwe, Martin Mutyaba, Solomon Segawa, Adonia Sebulime, Peter Lukwata, Rita Tukahirwa, David Ssimbwa, Hilda Mwesigwa and Margaret Nabatanzi. Thank you all for making this work enjoyable.

Lastly I would like to thank my wonderful family and friends. A special thank you to my parents, Mr and Mrs Richard and Juliet Sekiwano, for their enduring support and my lovely four year old daughter Elizabeth, for being on my team, “we have done this together”. Thank you Elizabeth, for all the beautiful pictures you have drawn of me although you have not done my hair any justice, those always reminded me that the thesis is not my life’s greatest achievement.

5 LIST OF PAPERS

This thesis is based on the following published papers, that will be cited as such;

I. **Nsangi A**, Semakula, D, Rosenbaum SE, Oxman AD, Oxman M, Morelli A, Austvoll-Dahlgren A, Kaseje M, Mugisha M, Uwitonze A, Glenton C, Lewin A, Fretheim A, Sewankambo NK.*Development of the informed health choices resources in four countries to teach primary school children to assess claims about treatment effects: a qualitative study employing a user-centred approach***.** Pilot Feasibility Stud **6,** 18 (2020). https://doi.org/10.1186/s40814-020-00565-6

II. **Nsangi A**, Semakula D, Oxman AD, Austvoll-Dahlgren A, Oxman M, Rosenbaum ES, Morelli A, Glenton C, Lewin S, Kaseje M, Chalmers I, Fretheim A, Ding Y, Sewankambo NK. *Effects of the Informed Health Choices primary school intervention on the ability of children in Uganda to assess the reliability of claims about treatment effects: a cluster-randomised controlled trial.* Lancet 2017; doi.org/10.1016/ S0140-6736(17)31226-6

III. **Nsangi A**, Semakula D, Glenton C, Lewin S, Oxman AD, Oxman M, Rosenbaum S, Austvoll-Dahlgren A, Nyirazinyoye L, Kaseje M, Rose CJ, Fretheim A, Sewankambo NK. *Informed Health Choices intervention to teach primary school children in low-income countries to assess claims about treatment effects: process evaluation.* BMJ Open 2019;9:e030787. doi: 10.1136/bmjopen-2019-030787

IV. **Nsangi A**, Semakula D, Oxman AD, Austvoll-Dahlgren A, Oxman M, Rosenbaum SE, Morelli A, Glenton C, Lewin A, Kaseje M, Chalmers I, Ding Y, Fretheim A, Sewankambo NK. *Effects of the Informed Health Choices primary school intervention on the ability of children in Uganda to assess the reliability of claims about treatment effects, 1-year follow-up: a cluster-randomised trial.* Trials 21**,** 27 (2020) doi:10.1186/s13063-019-3960-9

Other relevant publications contributed to by the author as part of the Informed Healthcare Choices Project

Nsangi A, Semakula D, Oxman AD, Sewankambo N. *Teaching children in low-income countries to assess claims about treatment effects: prioritization of Key Concepts*. Journal of Evidence-Based Medicine 2015; 8(4):173-80.

Austvoll-Dahlgren A, **Nsangi A**, Semakula D. *Interventions and assessment tools addressing key concepts people need to know to appraise claims about treatment effects: a systematic mapping review.* Systematic Reviews. 2016;5:215. doi:10.1186/s13643-016-0389-z.

Semakula D, **Nsangi A**, Oxman AD, Sewankambo N. *Priority setting for resources to improve information about claims of treatment effects in the mass media.* Journal of Evidence-Based Medicine 2015; 8(2):84-90.

Semakula D, **Nsangi A**, Oxman M, Rosenbaum S, Oxman AD, Austvoll-Dahlgren A, Glenton C, Lewin S, Kaseje M, Morelli A, Fretheim A, Sewankambo NK. *Development of mass media resources to improve the ability of parents of primary school children in Uganda to assess the trustworthiness of claims about the effects of treatments: a human-centred design approach. Pilot Feasibility Stud* **5,** 155 (2019). https://doi.org/10.1186/s40814-019-0540-4

Semakula D**, Nsangi A**, Oxman AD, Oxman M, Austvoll-Dahlgren A, Rosenbaum S, Morelli A, Glenton C, Lewin S, Kaseje M, Chalmers I, Fretheim A, Ding KY, Sewankambo N. *Effects*

*of the Informed Health Choices podcast on the ability of parents of primary school children in Uganda to assess claims about treatment effects: a randomised controlled trial***.** Lancet 2017; doi.org/10.1016/S0140-6736(17)31225-4

Semakula D, **Nsangi A**, Oxman AD, Oxman M, Austvoll-Dahlgren A, Rosenbaum S, Morelli A, Glenton C, Lewin S, Nyirazinyoye L, Kaseje M, Chalmers I, Fretheim A, Rose CJ, Sewankambo NK. *Effects of the Informed Health Choices podcast on the ability of parents of primary school children in Uganda to assess the trustworthiness of claims about treatment effects: one-year follow up of a randomised trial. Trials* 2020;21(1):187.

Semakula D, **Nsangi A**, Oxman A*,*Glenton C, Lewin S, Rosenbaum S, Oxman M, Kaseje M, Austvoll-Dahlgren A, Rose CJ, Fretheim A, Sewankambo NK.
Informed Health Choices media intervention for improving people's ability to critically appraise the trustworthiness of claims about treatment effects: a mixed-methods process evaluation of a randomised trial in Uganda
BMJ Open 2019;**9:**e031510. doi: 10.1136/bmjopen-2019-031510

Austvoll-Dahlgren A, Semakula D, **Nsangi A,** Oxman AD, Chalmers I, Rosenbaum S, Guttersrud Ø, and the IHC Group. *Measuring ability to assess claims about treatment effects: the development of the "Claim Evaluation Tools".* BMJ Open 2016;6:e013184. doi:10.1136/bmjopen-2016-013184.

Austvoll-Dahlgren A, Guttersrud Ø, **Nsangi A**, Semakula D, Oxman AD, The IHC group. *Measuring ability to assess claims about treatment effects: A latent trait analysis of the Claim Evaluation Tools using Rasch modelling.* BMJ Open 2017;0:e013185. doi:10.1136/ bmjopen-2016-013185.

6 LIST OF ABBREVIATIONS

IHC	Informed Health Choices
HRH	Human Resources for Health
SIHCLIC	Supporting Informed Healthcare Choices in Low Income Countries
SURE	Supporting Use of Research Evidence
WHO	World Health Organisation

7 PROLOGUE

The problem

I grew up in Sub-Saharan Africa during the height of the HIV/AIDS pandemic, which was accompanied by an explosion of unsubstantiated claims about the novel HIV virus, such as "*Using clay soil blessed by a local prophetess as a treatment for AIDS in the early 1990s in Uganda*". Unknowingly, my generation was shielded from many such claims, due to limited access to information, as well as the stigma that surrounded the illness.

Today, amidst the COVID-19 pandemic, we have unfiltered, virtually endless access to information online. Children, young people and adults alike are bombarded with both reliable and unreliable health claims from mass media, social media, friends, family, healthcare providers and public figures such as celebrities and even presidents.

Many of the claims are about benefits and harms of treatments (defined here as any action intended to maintain or improve the health of individuals or communities), including drugs, surgeries, lifestyle changes, diet, exercise, "alternative" medicine, public health and environmental interventions and changes in healthcare finance, delivery and governance.

People's inability to assess the reliability of health claims leads to misplaced trust in unreliable claims, which leads to harm and waste. This disproportionately affects low-income countries, with already limited resources.

The Informed Health Choices project

In 2012, as a freshly qualified graduate student, I was volunteering with the Supporting Use of Research Evidence (SURE) project led by Professor Nelson Sewankambo at Makerere University College of Health Sciences. This exposed me to the inner workings of policymakers' decision-making processes, and laid bare the lack of timely, relevant and "digestible" evidence that a policymaker could easily access, read, understand and interpret in order to make informed decisions.

I was approached by Professor Sewankambo about the Supporting Informed Health Choices in Low-Income Countries (SIHCLIC) project, which was related to SURE, only this time targeting laypeople, including children. I was fortunate to be selected as one of two PhD students for the project, which commenced January 2013.

SIHCLIC was later named the Informed Health Choices (IHC) project and included three main activities aimed at addressing the above-described problem;

1. identifying and organising concepts that people need to understand and apply when assessing claims about treatment effects and making health choices;
2. designing learning resources to help people master these concepts; and
3. evaluating the effects of the learning resources.

The project had two target populations: the general public (adults with no research background) and children of primary school age. My PhD work would focus on the latter.

Teaching primary school children to assess claims about treatments leverages on the time they have available for learning, while adults who tend to have less time to learn, as well as more to unlearn. Teaching children critical thinking can help them question the individual choices they make, creating a sense of personal responsibility, which is especially important during a pandemic, like the one we now face.

Creating a framework and using it to map interventions and measurement instruments

IHC started out by creating a framework of essential concepts that people needed to understand and apply in order to assess claims about treatment effects and make informed choices: the IHC Key Concepts. The initial framework included 32 concepts(1).

We then conducted a systematic review to map interventions and measurement instruments addressing one or more IHC Key Concepts, and found a total of 415 eligible studies (2). Only four of the interventions addressed 10 or more of the concepts.

Most of instruments measured health literacy, and did not directly assess people's ability to apply any of the IHC Key Concepts, but focused on functional literacy (reading and numeracy skills) (3, 4). Some of the instruments included critical appraisal skills, but these tended to measure more general understanding of health information and medical terminology (5-7).

Designing learning resources

We developed primary school resources to teach 12 of the IHC Key Concepts to primary school children aged 10-12 years, and a podcast to teach their parents nine of the same Key Concepts. We used an iterative human-centred design approach (8), with cycles of idea generation and prototyping (including workshops with stakeholders), user testing (appendix 1 and 2), piloting in classrooms and feedback from teachers in our teachers' network in Uganda. This approach ensures that the end-users experience the learning resources as useful, credible and appropriate for their own context. Most of the fieldwork was done in Uganda, but we also piloted and user-tested prototypes in Kenya, Rwanda and Norway.

Developing an outcome measure

To assess the effects of the resources, we developed the Claim Evaluation Tools database (9), having not found an adequate, previously-developed tool that measures people's ability to assess health claims and make informed health choices (2). The Claim Evaluation Tools database is a bank of multiple-choice questions, each measuring the ability to apply one of the IHC Key Concepts. The development was informed by extensive feedback from teachers, members of the public, methodological experts, health professionals and children. We conducted Rasch analyses to validate the questions (10, 11). Questions from the database can be used by researchers to evaluate interventions or map skills within a population; by teachers to assess learning; and by individuals for self-assessment.

Evaluating effects of the learning resources

Using a test with items from the Claim Evaluation Tools, we evaluated the IHC primary school resources and the podcast for parents in cluster-randomized trials (12, 13). The trial of the primary school intervention included a representative sample of 120 eligible schools with over 10,000 grade-five children in Uganda. The primary outcomes were mean score and the proportion of children achieving a predetermined passing score (≥13 of 24 correct answers) (14).

The mean scores in the intervention and controls schools were 62% vs. 43% (adjusted mean difference 20%, 95% CI 17–23%; p<0·00001). The proportion of children achieving a passing score was 69% in the intervention schools (3967/5753) vs. 27% in the controls (1186/4430) (adjusted difference 50%, 95% CI 44–55%). In subgroup analyses, the intervention was found to be effective in children with different levels of reading skills, but more effective in children with better reading skills.

The teachers took the same test as the children at the end of the term, with most teachers in both the control and intervention groups (87% vs. 98%) obtaining a passing score (adjusted difference 11%, 95% CI 4-13). The teachers in the intervention group were more likely to obtain a predetermined score indicating mastery of the concepts (72% vs. 15%; adjusted difference 57%, 95% CI 37-70)

In the linked randomized trial of the mass media intervention, we recruited a total of 561 parents of children at 35 of the 120 schools participating in the primary school resources' trial. The mean score for the 561 parents (83%) who completed follow-up in the podcast and control group was 68% vs. 52% (adjusted mean difference 16%, 95% CI 13–19%; p<0·0001). The proportion of parents achieving a predetermined passing score (≥11 of 18 correct answers) (14) in the podcast and control group was 71% (203/288) vs. 38% (103/273) (adjusted difference 34%, 95% CI 26–41%; p<0·0001). In other words, both interventions were found to have large, short-term effects on the primary outcomes (appendix 3).

We measured outcomes again after one year to assess retention (15, 16). Follow-up data showed that the learning is retained for at least a year, and conducted process evaluations to explore barriers and facilitators for scaling up use of the learning- resources, potential adverse effects and potential benefits(17, 18).

Resources from the IHC project are currently being translated and tested in other low-, middle-, and high-income countries. Building on this work, we are currently designing and evaluating learning-resources for secondary school students in East Africa, which will also lend themselves to translation and adaptation for use in other settings.

My experience

Working on this project, I have seen first-hand the damaging effects of an inability to appraise health information, but also that we can empower children as young as 10-year-olds with the necessary skills. For example, one of the children told us: “After learning the lesson about treatments that have been used for a long time, I went home and asked my mother to stop rubbing raw onions on my little brother’s nose.” The raw onion was used each time his brother had an epileptic convulsion.

Having the opportunity to help children learn to appraise health information, in particular claims about treatment effects, is one I will forever relish.

8 GENERAL INTRODUCTION

8.1 THE PROBLEM STATEMENT

Both reliable and unreliable health information about treatment effects is easily accessible to the general public through mass media, the internet, word of mouth and expert opinion. We encounter claims about the effects of treatments on a daily basis. Some claims are outright bogus, while others appear to be scientifically sound but are often based on poorly designed studies, expert opinions alone, and myths (19, 20).

Diverse evidence suggests that people's ability to appraise information is generally poor (21-26). For example, people appear to find anecdotal evidence more reliable, accessible and easily understood, compared to information derived from research evidence (27, 28). Surveys in the UK suggest that only about one third of the public trusts evidence from medical research, while about two thirds trust the experiences of friends and family (29). One possible explanation of this is that people may rely on "social proof" (30). This occurs typically in ambiguous social situations where people are unable to determine an appropriate course of action, and therefore do what others have done before, whether the actions of others were appropriate or not.

Social proof is one type of "conformity" where people are meant to believe that others' interpretation of an ambiguous situation is more accurate than their own (31, 32). Another possible explanation is that people have come to mistrust research because of conflicting and unreliable reporting in the mass media (33, 34).

Claims about treatment effects should be assessed based on systematic reviews of fair comparisons, in order to make informed decisions (33, 35). To recognise this, and to avoid being misled by bogus claims, people need education about key concepts relevant to assessing the effects of health interventions and interpreting research results. "Concepts" in this context are defined as criteria; i.e. standards for judgement or principles for assessing something. They are issues worthy of attention or consideration in assessing treatment claims and the evidence supporting those claims, and when making choices (36).

The list of Informed Health Choices (IHC) Key Concepts was originally developed to serve as a syllabus for identifying resources to help people understand and apply such concepts.

The IHC Key Concepts are only a starting point for teachers, researchers, and others to identify and develop resources to help people understand and apply the concepts (36).

Inadequate understanding of research evidence and intervention effects can increase or diminish the willingness of individuals to seek effective healthcare or participate in research. It can falsely raise expectations, dash hopes, and cause unnecessary alarm. This can result in decisions based on unreliable, incomplete, and sometimes harmful information (27).

Consequences of decisions that are based on unreliable claims include overuse of interventions that are ineffective and underuse of interventions that are effective. This in turn results in unneccesary suffering and waste of resources. Consequently, it is important to equip people with critical appraisal skills (37, 38).

The ability to assess if a claim is based on the best available evidence is one of the initial steps in making informed decisions (33) for patients, healthcare providers and policymakers. Studies suggest that patients can play a significant role in promoting evidence-based practice, and improvements in patient safety and the quality of care, if they are equipped to make informed health choices by actively participating in their care; for example, by asking clinicians relevant questions and holding policy makers accountable for their decisions (39-41).

The skills needed for critical appraisal of treatment claims are a subset of critical thinking skills. Critical thinking can be defined as “reasonable, reflective thinking focused on deciding what to believe or do” (42). It is an essential set of life skills relevant to decision making in many contexts.Thinking critically about what to believe and do about health is relevant and important to everyone, including children. According to the World Health Organization, good health is central to human happiness and well-being as it makes an important contribution to economic progress, as healthy populations live longer, are more productive, and save more (43). Because health is important and relevant to everyone, teaching critical thinking about health claims and choices is also a good way to teach concepts that can be generalised to claims about other types of interventions and other types of decisions (44).

Promoting and developing students’ critical thinking abilities is a core responsibility of the education system (45). Yet, many education systems still emphasize rote learning through memorisation, recital and repetition of facts rather than teaching thinking skills.

Critical thinking skills are included as a key competency in many frameworks for national curriculums for primary and secondary school in many countries (28). However, national assessments of student achievement too often focus on lower order thinking skills and learning instead of higher order thinking skills, reasoning and authentic performance (46). According to Blooms' taxonomy, "lower order thinking skills" include, remembering, understanding and applying, while "higher order thinking skills", include analyzing, evaluating and creating (47).

In Uganda, it is especially important to ensure that young people are enabled to recognise treatment claims, assess their trustworthiness, and make informed decisions to refrain from squandering meagre resources. Uganda is a youthful population with about 48% of it's population 14 years-old or younger (48). With a large proportion of the Ugandan population in the school going age bracket, there is an opportunity to teach them critical thinking skills while they are still in the education system.

In addition to obtaining health information from personal sources like teachers, healthcare professionals, parents, peers, and media sources such as newspapers and billboards, health information is now also easily available to Ugandan youth online. According to data from the Internet World Stats (2018), Uganda's internet usage has been on the rise since the start of the decade, with increased adoption of smart phones. It is estimated that about 35% of Africans on the continent have access to the internet as of 31st December 2017, with 31.2% of Ugandans having varied access to the internet compared to 0.1% of the population in 2000 (49). Although internet access is still out of reach for many Ugandans, there is a steady rise in the number of people that are accessing the internet, including online health information. A recent systematic review by Park and colleague (50) (2018) states that only 20% of African students reported spending an average of over 2 hours per day online compared to 42% and 40% of Chinese and American students respectively (51). While internet access can increase availabilty of reliable health information, it also increases exposure to unreliable claims about the effects of treatments(19, 52).

8.2 EXISTING RESEARCH EVIDENCE

What we know from previous research studies to address the problem

Efforts to measure and improve people's ability to assess claims about treatment effects have been focussed primarily on healthcare professionals (clinicians, nurses and medical students), usually targeting particular concepts like “understanding of risks” or “comparing like with like” (2). Efforts targeted at patients and the public have more often focused on health and functional health literacy (2). Health literacy is an evolving concept which is defined as people’s ability “to access, understand, appraise, and apply health information in order to make judgments and take decisions in everyday life concerning healthcare, disease prevention and health promotion” (53). Functional health literacy focuses on the “ability to read, write and perform basic mathematical calculations”(54, 55). Whilst research has documented that poor health and functional literacy is widespread and negatively impacts on people’s health (53, 56) , there has been a paucity of research focusing on the ability to think critically about health information and interpret research findings evaluating interventions (critical health literacy) (57, 58).

Evaluations of educational interventions that focus on critical appraisal of health claims have typically targeted health care professionals. Only a few studies have targeted children or adolescents (37, 59), and until recently, none of those have been in low or middle-income countries (37, 60, 61).

In a recent overview that included six sytematic reviews, totaling 227 studies on the effects of education interventions in low-income countries, none addressed critical thinking broadly, critical thinking within health literacy, or science literacy (62). Systematic reviews of studies teaching critical appraisal skills to children have not found any strategies for teaching such skills to primary school children that have been evaluated in either low or high income settings (37, 63).

In addition, existing measurement tools for measuring critical appraisal skills have predominantly been developed for health professionals, and the evaluation of published studies (64). Whilst there is a large number of critical appraisal measurement tools to choose from for healthcare professionals, there is still no consensus regarding what items should be included in these tools (65). A systematic review of interventions and assessment tools addressing key concepts that people need to know to appraise claims about treatment effects found 415 studies that addressed one or more of the IHC Key Concepts (2). However, only four assessment tools included 10 or more IHC Key Concepts and none of these were targeted at patients or members of the public including children. Among the studies identified, the most commonly assessed concepts were:

- “treatments usually have beneficial and harmful effects” (273 studies),
- “treatment comparisons should be fair” (131 studies),

- “compare like with like” (117 studies), and
- “single studies can be misleading” (43 studies).

Most of the IHC Key Concepts were included in at least one study, with the exception of three concepts:
- “more is not necessarily better”,
- “avoiding unrealistic expectations”, and
- “average differences can be misleading”.

These measurement tools were frequently designed for evaluating a specific intervention. Few of the outcome measures used in the studies of educational interventions have been validated (66).

9 JUSTIFICATION

9.1 WHY THE FOCUS ON PRIMARY SCHOOL CHILDREN?

Teaching children critical thinking skills early in life can lay the foundation for future learning. It may also help to foster desirable dispositions (habits of the mind), such as questioning the basis for treatment claims, and help to hinder uncritical beliefs, which may be difficult to change as children grow older (67). Teaching critical thinking skills may also improve academic achievement (68-70).

Children aged 10-12 are capable of learning critical thinking skills (71), and in several countries teaching these basic skills is already part of the curriculum (72). Although it is still possible to teach adults critical thinking skills (13), it is likely to be easier if the ground has been laid early on in childhood. Adults have far less time to learn, and there is less expectation for them to learn. Children have time to learn while still at school and they are expected to do so. If school children are targeted, it makes it possible to reach a significant part of the population in low-income countries like Uganda, which tend to have notably young populations and a large number of school dropouts before completing primary level (73-75).

Health and education are high priorities in most countries and are interdependent Sustainable Development Goals (76). The education system plays an important role in preparing people to be informed healthcare consumers, practitioners, and policy makers (77, 78). For example, several studies (79-82) have found an association between a woman's level of education and her children's health at birth (such as a reduction in the likelihood of low birth weight and premature births). Such findings suggest that education of women can lead to a substantial non-pecuniary benefit in terms of the health of their offspring, especially in developing countries. Thus, teaching primary school children to think critically can potentially help to achieve Sustainable Development Goals for both health and education.

9.2 THE UGANDAN CONTEXT

The Health sector

Uganda is ranked at 143 out of 169 countries surveyed for the 2010 Human Development Index (83). Despite achieving a 5.8% gross domestic product (GDP) growth in 2016, a third of the country still lives on less than one US Dollar a day. About 85% of the current population lives in rural areas. In urban areas, 60.1% of the urban population resides in slums.

Life expectancy has slowly been increasing from 40 years for males and 42 years for females in the mid-nineties during the height of the HIV epidemic, to its current level of 57 years for males and 61 years for females (84). The infant mortality rate is 66 per 1,000 live births compared to a global average of 43 infants per 1000 live births (85). Uganda's maternal mortality ratio is one of the highest in the world, standing at 342 per 100,000 live births, compared to a global average of 260 per 100,000 live births. Only about 42% of births are attended by skilled health personnel (85).

Uganda has one of the worst doctor-patient ratios in the world, according to the World Health Organization (WHO) (85), with a doctor-patient ratio estimated at 1:24,725 and the nurse-patient ratio of 1:11,000. The WHO's current doctor-patient recommendation is 1:1000.

There are significant disparities in access to health care due to a multitude of factors including infrastructure limitations, insufficient health human resource, poor financing and logistics, and poor access to information on available and efficient healthcare interventions.

The WHO defines human resources for health (HRH) crisis countries, as severely constrained, with 1 health professional per 435 people (2.3 per 1000) (86). At 1 health professional per 700, Uganda is well below the WHO minimum standard.

Living in a resource-limited setting with one of the worst doctor-patient ratios in the world, it is perhaps not surprising that many Ugandans are lured by alternative remedies, usually with little or no evidence to support claims about their effectiveness.

The Education sector

The current Ugandan primary school thematic curriculum that was adopted in 2007 is based on three main principles (87):

i) rapid development of literacy, numeracy, and life skills at lower primary (grade 1 to grade 3);

ii) the treatment of concepts holistically, under themes of immediate meaning and relevance to the learner (life skills); and

iii) the presentation of learning experiences in languages in which the learners are already proficient (multi-lingual).

Nonetheless, the current curriculum covers almost the same areas that existed in the old primary school curriculum. Recent educational/ curriculum reform efforts in Uganda have lacked a realistic and substantial investment in physical resources such as infrastructure (classrooms and toilet facilities), failed to address systemic issues (high student-teacher ratios, lack of quality textbooks) and human capital (teacher training, remuneration, and retention) (88). The disparity between the expectations resulting from the reforms and the reality in the classroom may be due to the adoption and use of reforms unsupported by evidence.

Although the current curriculum strives to adopt a ‘child-centred approach’ by putting the child’s interests, experience, and needs at its centre, it still emphasizes the acquisition of facts in various subjects with emphasis on recall and other lower cognitive skills.

Teachers are expected to assess their students routinely for diagnostic and remedial purposes. This focuses on identifying students’ specific areas of weakness in a given course of instruction and then suggesting remedial measures. It is believed that frequent assessment facilitates timely feedback and corrective action. This is intended to enable teachers to identify problems and address them as soon as they are identified, and to tailor solutions to both high and low achievers (87). This expectation may be difficult to meet, as teachers grapple with large class sizes on a daily basis (88), with an average teacher-student ratio of 1:69 (89).

The national primary school curriculum currently lacks a syllabus for critical thinking as a skill and no previous effort has been made to identify effective strategies of teaching critical thinking in the Ugandan context. While critical thinking is now one of the main learning objectives of the newly introduced lower secondary school curriculum starting in the academic year 2020, there is currently no consensus on teaching and evaluation strategies.

10 OVERVIEW OF THE RESEARCH PROJECT

10.1 Overview of the Informed Health Choices Project

The Supporting Informed Health Choices in Low-Income Countries (SIHCLIC) project was funded by the Research Council of Norway. We have since subsequently referred to the project as the Informed Health Choices (IHC) Project. The objectives of the project were to develop and evaluate two strategies for improving health literacy;

i) learning resources for primary school children
ii) mass media resources for the general public.

The IHC research team included researchers from Kenya, Norway, Rwanda, Uganda, and the United Kingdom. Members of this multidisciplinary research team had experience in research methodology; health literacy; and developing and evaluating strategies for communicating research evidence and teaching evidence-based practice to health professionals, patients, journalists, and policy makers.

The work reported in this thesis focusses on the first objective: developing and evaluating learning resources for primary school children. Prior to developing and evaluating those resources we needed to identify the key concepts that should be taught to and learned by primary school children. We also needed to develop an outcome measure. I describe that work here, before discussing the research that is included in the thesis. We engaged stakeholders throughout the project, and I begin by presenting the main ways in which we did this.

10.2 Stakeholder Engagement

Involving teachers and children in the design of the intervention may have made an essential contribution to its effectiveness as described in Paper I (development of the intervention) and documented in Paper III (process evaluation) of this thesis. In addition, we engaged stakeholders in this project from the beginning through three groups.

International Advisory Panel

Using purposive sampling and suggestions from members of the Informed Health Choices Group, we compiled a list of people who have relevant experience and expertise. This included people with expertise in education, communication, health literacy, and evidence-based medicine. We invited people from different countries, including low and middle-income countries, to be members of the International Advisory Panel.

The International Advisory Panel consisted of a total of 29 members, and their main contribution was to provide general feedback on a draft list of IHC key concepts, which provided a framework for developing the intervention, as reported in the first paper in this thesis. Although we had initially intended to engage the advisory group in the design of the intervention, this was only done to a very limited extent.

National Advisory Panel

Using a similar approach as for the International Advisory Panel, we identified a group of stakeholders that we invited to consultation meetings in Uganda. The consultation meetings were used to inform key stakeholders about the work we were doing and gather support for the project in the region.

The National Advisory Panel that we established included 15 members with a diverse range of backgrounds. This group included policy makers at the ministries of health, education, and social development; media house representatives; primary school teachers; and representatives of civil society.

The consultation meetings were important for establishing a working relationship between the researchers and the policy makers. They also provided input on effective ways of engaging teachers, primary school children, and the community in the project.

Teachers' Network

With input from policy makers at the Ministry of Education, the research team identified two districts, Kampala and Wakiso, that were close to the research site and had diverse characteristics of primary schools in terms of ownership (government funded or privately owned) and location (urban, semi-urban, or rural). We established a teachers' network that included 24 volunteers that were selected from a list of schools supplied by the educational officers in the two districts.

We aimed to get a small sample of primary school teachers, with varying backgrounds (education, qualifications and experience). We recruited roughly 12 teachers from each of the two districts. Using a multistage stratified sampling method in each district, primary schools were identified from both government and privately funded schools (first stage). The two categories (government and private) were further divided into rural and urban schools (second stage) (90). We then purposively selected schools from each identified category to take part in the research.

The education authority in each district provided invitation letters introducing the research team to the selected schools. The research team met with the head teachers for each selected school to give them a brief overview of the project, before asking them to identify a suitable teacher to work with the project as a member of the teachers' network on a voluntary basis.

The research team followed up with letters formally inviting the teachers to join the teachers' network. A written confirmation of acceptance was obtained from each teacher. The teachers were invited to attend a three-day meeting where the IHC key concepts identified prior to the meeting were prioritized for primary school children (91). The teachers network was also engaged in the development of the intervention, as described in paper I of this thesis.

10.3 Initial identification of the IHC Key Concepts

We identified an initial set of concepts by extracting them from Testing Treatments (35), a book that was written to promote more critical public assessment of claims about the effects of treatments. We then reviewed other material written for the general public, (21, 35, 92), for journalists (93-95), and for health professionals(93, 94, 96). Through a series of meetings between the research team and the international advisory panel, the list of concepts was iteratively discussed, revised, and grouped in a coherent and logical way.

Our focus was on concepts that children or adults with a primary school education and basic reading and numeracy skills could potentially learn to understand and apply. At the same time, the list was intended to be universally relevant (to all age groups), and we included some concepts that would be difficult for primary school children to learn. We used plain language to describe and explain each concept. However, the list of concepts was not intended to be used as a learning resource. It was a syllabus or framework for developing learning resources to teach the concepts. The initial list included 32 concepts. The list is currently an evolving document which is revised yearly (36, 97).

Having developed an inventory of key concepts for learning resources that are needed for assessing treatment claims, we then needed to determine which of those concepts could be learned by primary school children. To do this, we organised three consecutive day-long meetings with the teachers' network. At these meetings we first introduced and explained the concepts to the teachers.

It was clear at the meetings that the concepts were new to many of the teachers. We drew on local and easily understood examples of claims about treatment effects and encouraged the teachers to think of other relevant examples that could easily be understood by a primary school child. We then used a modified Delphi Technique (98, 99) to prioritise key concepts based on pre-set criteria which included,

(i) relevance of concepts for children,
(ii) ease of comprehension of concepts for children,
(iii) potential for developing resources to teach the children,
(iv) whether the resources once developed would have an impact on children's ability to assess claims.

The teachers agreed that all the concepts were relevant, but that some concepts were more complex than others for the children to understand (91). This exercise informed subsequent decisions by the research group about which key concepts we would try to teach the children.

10.4 Development of the outcome measure

The CLAIM Evaluation Tools Database was developed by the project team to objectively measure participants' ability to apply the key concepts. The database is a flexible battery of multiple-choice questions from which researchers or teachers can select questions for the concepts they are interested in.

The database was developed over a period of three years using both qualitative and quantitative methods. The process used to develop the database and validate the outcome measure that was used in evaluating the primary school learning resources included:

i) a systematic review of exisiting measurement instruments,

ii) development of the database,
iii) validating the tests used in the evaluation; and
iv) establishing cut-offs for a passing score and mastery of the concepts included in the learning resources.

Systematic review of existing measurement instruments

We conducted a systematic review mapping interventions and assessment tools addressing the key concepts people need to know to appraise claims about treatment effects (2). The review was conducted to identify suitable measurement tools to be used in the IHC trials, alternatively to inform the development of an evaluation tool, if we could not find a suitable tool. The review included 415 studies of interventions and assessment tools. Our list of key concepts (1) was used as a framework for the review and guided the identification of relevant interventions and assessment tools. The review included studies aimed at patients, the general public, and healthcare professionals.

The studies retrieved during the review process were categorised in four broad areas:
i) risk communication and decision aids studies, mainly targeting patients;
ii) understanding of trials, mainly targeting trial participants in order to improve recruitment and retention;
iii) evidence based and critical appraisal studies, predominantly targeting health care professionals; and
iv) science education studies, mainly targeting school children.

The systematic review concluded, that a significant number of the key concepts (1) were included in at least one intervention. However, most of the interventions targeted only a handful of the key concepts, ranging from a couple of concepts to less than half of the concepts.

The instruments identified in the review targeted healthcare professionals, medical students, health communicators, patients (children and adults), and research participants. None targeted healthy children or ordinary people with no research training (2). We did not find an instrument that addressed all the IHC Key Concepts or that was suitable as an outcome measurement tool for the IHC school trial in Uganda. Findings from the systematic review informed the development of a database of multiple-choice items that could be used as an outcome measurement tool.

Development of the database

The Claim Evaluation Tools Database is a set of flexible items to measure people's ability to assess claims about treatments. It was developed with input from children, teachers, and methodology experts in Uganda, Kenya, Rwanda, Norway, United Kingdom, and Australia. We applied purposive sampling while selecting study participants. We included teachers of evidence-based medicine, individuals with research methodology training, patients and members of the public in both low and high-income countries.

The development of the Claim Evaluation Tools included:

i) determing the scope of the database and development of items,

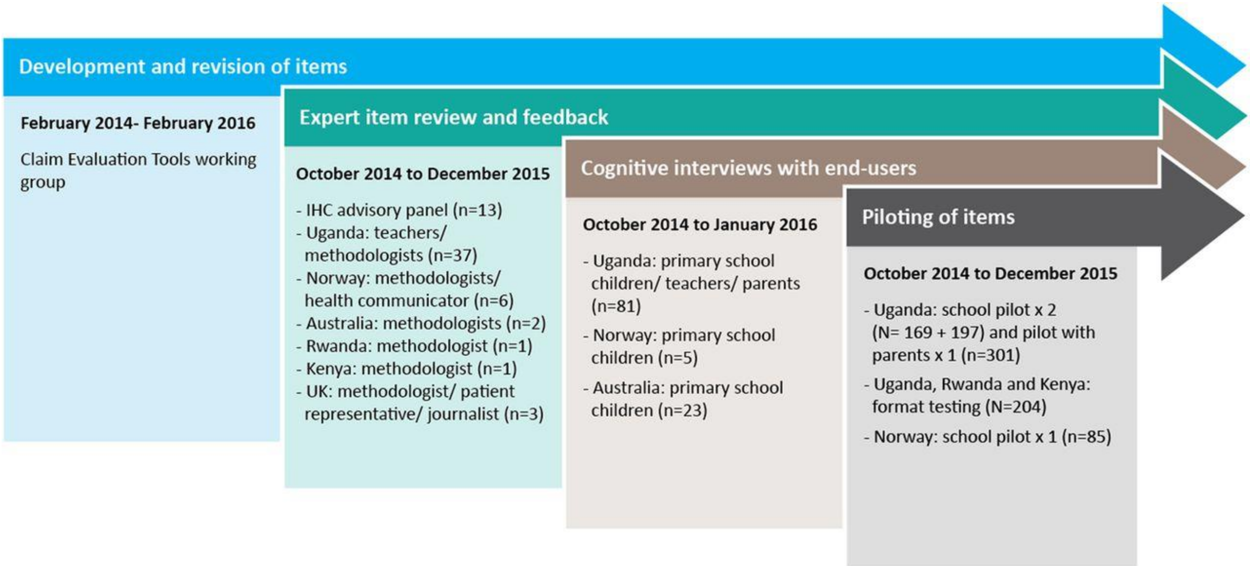
ii) expert item review and feedback (n=63),

iii) cognitive interviews with children and adult end users (n=109), and

iv) piloting and administrative tests (n=956).

An overview of the development process is shown in Figure 1.

Figure 1: Overview of the CLAIM Evaluation Tools development process



The resulting Claim Evaluation Tools Database is a flexible resource which includes a battery of multiple-choice items, with each item beginning with a scenario intended to be relevant across contexts and which can be used for younger children (10 and above), adult members of the public and healthcare professionals as well as adapted by researchers, teachers and other users to design measurement instruments tailored to their own use.

Research methodology experts judged the items to have face validity, while end users, including teachers, found them to be relevant and acceptable for their particular settings. Based on feedback from study participants (research methodological experts and end users), items were revised, texts simplified, simple explanation terms were added, and we redesigned formats and instructions.

The Claim Evaluation Tools are freely available for non-commercial use (on request) through the IHC website.

Validation of the tests used in the evaluation

We assessed items from the CLAIM Evaluation Tools database using Rasch analysis to develop an outcome measure for our intervention trial. The main objective was to assess the validity and reliability of 88 items addressing 22 key concepts. Rasch analysis is a form of psychometric testing relying on Item Response Theory (100). It is used to develop outcome measures that are valid and reliable.

Rasch analysis provides a basis for building and revising items in question banks. Items diagnosed as misfits to the Rasch model can be deleted or revised to improve model fit (100). Rasch analysis is also used to validate outcome measures' internal construct validity, multidimensionality, invariance of the items (item-person interaction), item bias (differential item functioning), and the degree to which scoring and summing up across items is defensible.

We conducted two Rasch Analyses (10, 11, 101). Data collection mostly took place in Uganda, where the items from the CLAIM Evaluation Tools database were intended to be used as the primary outcome measure for the school trial. The sample used to evaluate the items during the two Rasch analyses, included both children and adults in order to explore item bias differential item functioning associated with age, as well as people with different backgrounds (healthcare professionals, patients, healthy adults and children who had and hadn't participated in the piloting of the IHC resources). A small sample of children who had participated in piloting the primary school resources at an international school in Norway were included to provide an indication of fit of the Rasch model, information on difficulty and differential item functioning in an international population compared to a low-income setting.

Four sets of multiple-choice items were administered to 1114 people in Norway and Uganda (685 children and 429 adults; including 171 healthcare professionals).

Findings from the first Rasch analysis were used to revise items in the CLAIM Evaluation Tools Database. Although most of the items conformed well to the Rasch model, some needed revision. All four sets of items that were tested, however, had satisfactory reliability overall. In addition, the items had a high level of difficulty. After revision we concluded that most of the items were suitable for use in an outcome measure for evaluating the ability of children and adults to assess treatment claims.

The objectives of the second Rasch analysis were to evaluate two sets of items for use in the IHC primary school trial (12) and a trial of an intervention directed at parents of school children (13). Each test included 26 multiple-choice questions for the 13 key concepts that we would teach in the primary school and podcast trials. The two tests were administered as oral tests in Luganda and written tests in English to a total of 1617 children and adults in Uganda. Based on the findings from this study, we chose the items with the best fit to the Rasch model for use as the outcome measure in the trial.

Determining cut-offs for passing and mastery scores

We recruited eight independent judges to determine the cut-off for a passing score (having at least a borderline ability to apply the concepts) and mastery of the concepts included in the learning resources (14). We used a combination of Nedelsky's and Angoff's methods (102, 103), two approaches that have been used for setting standards for performance on educational and licensing tests.

The eight judges had backgrounds in research methodology, evidence-based medicine, teaching, and education research. They independently assessed the probability for someone with borderline ability to apply the concepts answering each question correctly. They did the same for someone that had mastered the concepts. These probabilities were summarised to determine a cut-off for passing and for mastery. A summary of their judgements was prepared and discussed, and a consensus was reached about a final passing score for the 12 key concepts included in the IHC primary school resources (with two questions for each concept).

11 STUDY OBJECTIVES

The objectives of the research in this thesis were to:

- Design and user test learning resources to enable primary school children to understand and apply key concepts for critically assessing treatment claims (Paper I)
- Evaluate the effectiveness of the primary school learning resources in a randomised trial (Paper II)
- Identify factors affecting the implementation, impact, and scaling up of the intervention; and any potential adverse and beneficial effects (Paper III)
- Evaluate the effectiveness of the primary school learning resources one year post intervention (Paper IV)

12 METHODS

12.1 Development of the intervention (IHC primary school resources) -Paper I

The Objective for paper I was to design and user test learning resources to enable primary school children to understand and apply key concepts.

Participants and setting

We recruited schools that were geographically accessible to the research team, were willing to make time in their schedules for piloting and user testing, and used English as a language of instruction. We requested head teachers of the participating schools to identify and introduce the research team to science teachers and children who would be willing to pilot the resources. Participants in the user testing interviews were recruited through purposeful sampling, to include grade five children (10 to 12-year olds) and their teachers.

Although most of the piloting and user testing took place in Uganda, our initial plan was to create resources that may be usable in other settings. In order to ensure usability of the resources in other contexts, we also carried out user testing interviews and pilots in Kenya, Rwanda, and Norway.

The development process

The development of the resources to teach primary school children to assess treatment claims was a collaboration amongst researchers, information designers, teachers, and children in Norway, the United Kingdom, Uganda, Kenya and Rwanda. We employed a user-centred approach characterised by multiple iterative cycles of development (104-106).

The initial starting point for developing the learning resources was the list of 32 key concepts deemed important for people to understand in order to appraise claims about treatment effects (91). The relevance of the various concepts to primary school children was assessed by the network of 24 Ugandan primary school teachers (91).

The process to develop teaching resources comprised of three main stages:

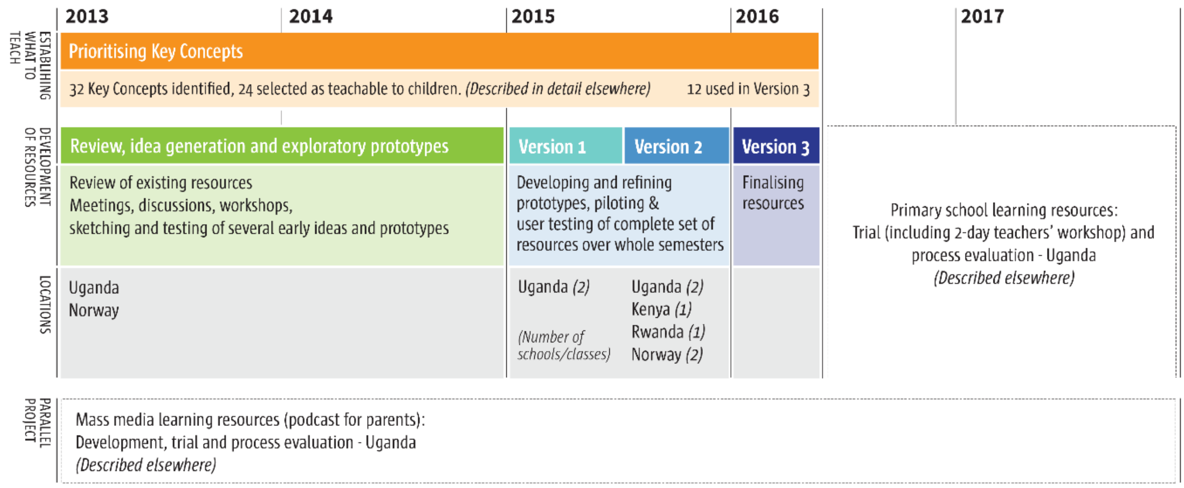
(i) idea generation and prototyping,

(ii) piloting and user testing, and

(iii) analysis and revisions.

An overview of the development is shown in Figure 2.

Figure 2: Timeline showing the development process of the primary school resources



The brainstorming workshops involved end-users - primary school teachers and children. "Brainstorming" is a creative technique that allows for gathering a list of spontaneous ideas without inhibitions to address a specific problem (107).

Several workshops were conducted both in Uganda and Norway (91). These resulted in ideas, contextual awareness, and prototypes. Ideas with the most potential were selected for further development, building on them to create new prototypes, which were used in the next phases of pilot and user testing. We used creative thinking to generate ideas prototypes. Creative thinking focusses on exploring as many ideas as possible (108), quite contrary to critical thinking which focuses on identifying the correct answer and eliminating incorrect options. Using both types of thinking was necessary to generate appropriate options for the intended resources.

Involvement of primary school teachers and children in all the development phases was necessary to ensure appropriateness and usability of the resources, as none of the research team belonged to the end user groups.

Members of the teachers' network collaborated closely with the research team during the brainstorming and prototyping workshops. Feedback was collected at workshops and school visits.

We interviewed end users (teachers and children) to explore their perceptions of the resources and their experiences using them (104). “User testing” comes from human computer interaction, which requires that effectiveness and efficiency of a product is measured in relation to an individuals’ satisfaction while using the product.

Rosenbaum’s modified version (109) of Peter Moville’s honeycomb framework (104), was used to develop interview guides for individual interviews and focus group discussions. We focussed on six facets of the users’ experiences: understandability, ease of use, desirability, usefulness, credibility, and identification (appropriateness) (104).

The earliest prototypes were piloted in workshops with teachers and children in Uganda and Norway. We used a non-participatory observation approach to facilitate user engagement. Once we had more complete prototypes (version 1 in Uganda, and version 2 in Uganda, Rwanda, Kenya, and Norway), we used video, still photographs and structured forms to record observations during ongoing lessons in classrooms, as teachers and children used the materials.

Analysis and revisions

Using a spreadsheet, we entered our observations from the pilots and feedback from user-testing interviews. The findings of these observations were independently coded by at least two researchers based on importance of the finding and its’ implications to the learning resources being developed.

12.2 Evaluation of the learning resources post intervention - Paper II and Paper IV

The objective for paper II was to evaluate the effectiveness of the primary school learning resourcesm, and for paper IV to evaluate the effectiveness of the primary school learning resources, one year post intervention.

Participants

We randomly selected four districts in central Uganda. Together with the educational officers of those districts, we generated a list of 2960 potential schools. Of those, 931 were excluded for various reasons such as prior participation in piloting the resources during the development process, inaccessibility due to long travel time, and schools without grade five.

Of the remaining schools, 170 were randomly selected and invited to recruitment meetings from April 11th to June 8th 2016. Of those invited, 120 agreed to take part in the trial. We included all grade five children in each of the participating schools.

We informed the head teachers about the purpose of the study during the initial recruitment meetings, before they consented on behalf of their school to participate in the study. After randomisation, the school heads were informed of the study arm to which their school was allocated. The consent form also included information about the outcome measure, stating that "it consist of multiple-choice questions that assess one's ability to apply the key concepts that people must understand and apply to assess treatment claims and to make informed health choices." We did not show the children or the teachers the test until they completed it at the end of the school term and a second time one year later. Children in both arms were informed that the test was being used to evaluate the IHC primary school resources when their teachers asked them to complete it.

Study design

The study was a cluster randomised trial with follow-up at the end of the term when the intervention was delivered, and again after one year. We randomly allocated schools to the intervention or the control group using a computer generated sequence with block sizes of four and six and equal allocation ratios within each block

In order to ensure equal distribution of the schools for school ownership (public or private) and geographical location (urban, semi-urban, or rural), an independent statistician generated six randomisation lists - one for each combination of the two variables. The statistician and his assistants labelled opaque envelops with the unique codes, inserted cards with the study group allocated to each code in the envelopes, and sealed them.

After obtaining consent to participate from the 120 school heads, two research assistants selected each school from the list of schools and identified the appropriate randomisation lists to be used for that particular school based on the school's location and ownership. The next available code from that list was assigned to the school before opening the corresponding envelope to determine whether the school was assigned to the intervention or control group. No changes to allocation were made during or after the randomisation process.

Interventions

The learning resources were designed to be used over one school term, (nine weeks) with one double period (80 minutes) per week, and one hour to complete the test at the end of the nine weeks. The intervention period was in the second term of the academic year, from June to August 2016.

All teachers in the intervention arm attended a two day introductory workshop, where we informed them about the study objectives, procedures, went through all the nine lessons in the primary school resources, and addressed any concerns that arose. Teachers from schools in the control arm were invited to a two hour introductory meeting to inform them about the study procedures and the outcome measurement tool. Control teachers were not introduced to the resources or invited to the introductory workshop.

None of the schools in the intervention arm continued using the resources between the school term when the intervention was delivered and the one year follow-up. We did not intervene in the control schools and the resources were not provided to those schools until after the one year follow-up. All schools in both arms of the trial administered the test at the end of the term when the intervention was delivered and again after one year, in August 2017.

Outcomes

The primary outcomes were:

- The mean test score (percent of correct answers)
- The proportion of children with a passing score

The secondary outcomes were:

- The proportion of children with a passing score for a sub-group of children who received an audio version of the test in Luganda (post intervention only)
- The proportion of children with a score indicating mastery of the concepts at the end of the intervention term and again one year later
- For each concept, the proportion of children who answered both questions correctly at the end of the intervention term and again one year later
- The children's intended behaviours and self-efficacy at the end of the intervention term and again one year later
- Self-reported behaviours (one year follow-up only)
- The children's attitudes towards science and school (post intervention only)

- Mean scores, passing scores, and mastery scores for the teachers, who took the same test as the children at the end of the intervention term and again one year later
- Mean attendance and mean scores on national exams for the intervention term and the following term (reported in one year follow-up study)

The teachers completed the test at the same time as the children. The test included 24 multiple-choice questions, two for each of the 12 concepts resulting from the selection process described in paper I. In addition to the 24 questions, two additional multiple-choice questions that were not covered in the primary school resources were included because the test used in these trials was also used in a linked randomised trial evaluating a podcast given to the parents of some of the children (13). Responses to the two extra questions were not included in the primary analyses. The test also included questions that assessed intended behaviours, self-efficacy and attitudes. Additionally there were four questions that assessed reading skills, and responses to those four questions were used as a covariate in an exploratory analysis.

An absolute (criterion referenced) standard was used as a cut-off for passing and mastery scores (14). To obtain a passing score, children had to answer as least 13 out of 24 questions correctly, to obtain a mastery score at least 20 out of 24.

Statistical analysis

The statisticians who analysed the data did not have prior information about which group was the intervention and which one was the control group when the primary analyses were done. For the analysis of primary and secondary outcomes, we used mixed models with a random effects term for clusters and the stratification variables modelled as fixed effects. We used generalised logistic regression for dichotomous outcomes and linear regression for continuous outcomes. All participants for whom there were data were included in the analyses.

12.3 Process Evaluation - Paper III

This was a multi-method study using qualitative and quantitative data. The objective of this paper was to identify factors affecting the implementation, impact, and scaling up of the intervention; and any potential adverse and beneficial effects.

Participants and setting

For qualitative data collection, we sampled six of the 60 schools that participated in the intervention arm of the trial. Purposive sampling was applied based on school location (rural, semi-urban, or urban) and school ownership (public or private), and variation in the extent to which the teachers implemented the intervention as intended. Within each school we included the head teacher, all grade five teachers who participated in the trial, two children and two parents. Individual children were sampled based on their performance on the school end of term assessments and how well they understood the lessons. Parents were sampled from both arms of the trial. We intended to include all district education officers in the districts where the trial took place.

Qualitative data analysis

We used a framework thematic analysis approach to guide data collection; familiarisation, coding, charting, and interpretation of the data; and analysis (110). Data were collected using lesson evaluation forms, structured observation forms, individual and focus group interviews.

We appraised the certainty of our findings using a modified version of the GRADE-CERQual approach (111).

Quantitative data analysis

Quantitative data were obtained from the lesson evaluation forms that all 85 teachers in the intervention arm completed. These included their assessments of the suitability of the teaching resources on a Likert scale of one to six (1=lowest, 6=highest).

13 ETHICS APPROVAL AND INFORMED CONSENT

We obtained ethics approval for all the four studies from the Uganda National Council of Science and Technology and Makerere University Institutional Review Board. We provided the head teacher of each school with information about the study and obtained written consent from them on behalf of their school to take part. We also informed the grade five teachers who participated in the development phase (user-testing and piloting of the materials), evaluation (trial and one year follow up study) and the process evaluation about the study procedures and practical impact on their practice before obtaining their written consent. We sought parental consent from the parents of the children who participated in the focus group discussions and individual interviews, and from the parents who participated in interviews. We informed all the children in the trial about the studies but we did not seek their parents' written consent or obtain individual assent from the children.

14 STUDY RESULTS

14.1 Development of the intervention (IHC primary school resources) -Paper I

We generated a large number of ideas for learning resources at workshops with the research team and the teachers' network. Many of these were learning games and other classroom activities. After prototyping some of these ideas, we identified two major problems. First, it was very difficult to organise games and activities in classes with 70 to 100 children and one teacher. This was especially difficult for games that required material to be distributed and games with which the children were not familiar. Second, because the key concepts were new to the teachers, they needed training or support to begin teaching the concepts while they were, to some extent, learning the concepts together with the children. Our solution to the first problem was to only use activities that could easily be managed in classrooms with a high student to teacher ratio and few resources. Our solution to the second problem was to create a textbook for the children, using a comic book story to introduce the concepts, and to provide the teachers with a guide. For each chapter there were exercises for the children to do individually, and a classroom activity. We tested a prototype of the textbook with just two chapters, and found this solution to be promising.

We then prepared a prototype that included all 24 key concepts that we had originally selected for primary school children. Pilot and user testing of two versions of a textbook that included all 24 concepts revealed that this was too many concepts to teach in a single school term. We also found that the normal school lesson (40 minutes) was too short to read the lesson aloud, allow time for individual children to do the exercises, and have time for the class activity in each lesson. In addition, teachers needed time to recap the previous lesson. We therefore decided to only include 12 of the 24 concepts selected in the textbook and accompanying resources for a single school term. In addition, we designed each lesson to be taught in a double lesson (80 minutes).

The resulting resources that we developed included the children's textbook and exercise book, a teachers' guide, a classroom poster of the key concepts, a set of activity cards for one of the chapters, and a song (Figure 3). These open access resources can be viewed or downloaded at the IHC website.

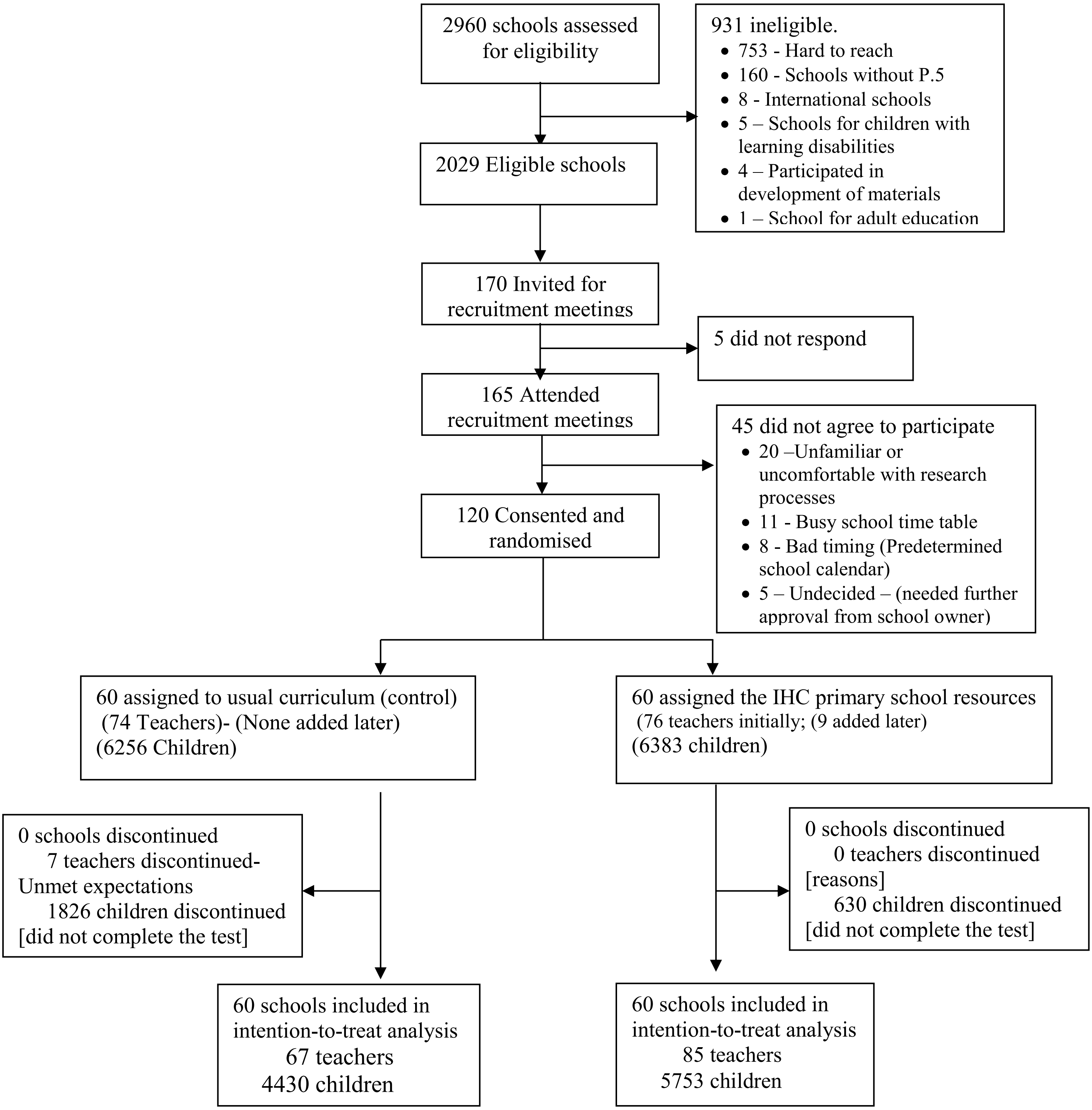
Figure 3: Pictures of the Informed Health Choices Primary School Resources

14.2 Effects of the learning resources - Paper II

Of the 170 schools invited to recruitment meetings, 120 agreed to participate in the trial. The most common reason for declining to participate in the trial was unfamiliarity with the research process and a busy school time table. The schools were randomised into two equal sized groups that proved to be similar with regards to their baseline characteristics. The trial profile is illustrated in Figure 4.

Although the timing of the lessons varied, all the intervention schools delivered all the nine lessons. A total of 10,183 children completed the evaluation test. In the intervention arm, 90% of the children completed the questionnaire compared to 71% in the control arm. This may have been the result of intervention teachers being more motivated to request that children stay at the end of the term to take the test than the control teachers.

Figure 4. Trial profile



Primary outcomes

The average score for children in the intervention schools was 62.4% compared to 43.1% in the control schools. The adjusted mean difference (based on regression analysis) was 20.0% (95% CI 17.3-22.7; p<0.00001).

In the intervention group, 69% of the children had a passing score (≥ 13 out of 24 correct answers) compared to 27% in the control group. The adjusted difference (based on the odds ratio from the logistic regression analysis) was 50% more children who passed (95% CI 44-55; p<0.00001) in the intervention than in the control group.

Secondary outcomes

- The average score in the intervention schools was 66.3% compared to 49.7% in the control schools for the proportion of children (1616 children) who completed the oral test (an audio version of the test in Luganda). The adjusted difference was 15.8% (95% CI 12.7-19.0).
- In the intervention schools, 19% of the children had a mastery score (≥ 20 out of 24 correct answers) compared to 1% in the control schools. The adjusted difference was 18.0% more children in the intervention schools who mastered the concepts (95% CI 18-18; p<0.00001).
- For each concept, the proportion of children who answered both questions correctly was higher in the intervention schools than in the control schools, including for the concept that was not taught to the intervention children (that a treatment outcome may be associated with a treatment, but not caused by the treatment).
- Children in the intervention schools were more likely to respond that they would find out what a claim was based on (adjusted difference 10.6%, 95% CI 6.2-14.7); find if a claim was based on a research study (10.8%, 6.3-15.1); and participate in a research study (7.8%, 3.7-11.9) compared with their counterparts in the control schools.
- Intervention children were more likely to consider it easy to assess whether a claim is based on research (adjusted difference 15.0%, 95% CI 10.9-19.0) compared to those in the control.
- Children in the intervention arm were less likely to consider it easy to assess how sure they could be about research results (adjusted difference -4.1%, 95% CI -1.0 to -7.3)

- In the intervention arm, 98% of the teachers had a passing score (≥ 13 out of 24 correct answers) on the same test that the children took at the end of term compared to 87% in the control (adjusted difference 11%, 95% CI 4-13)
- In the intervention group, 72% of the teachers achieved a mastery score (≥ 20 out of 24 correct answers) compared to 15% in the control (adjusted difference 57%, 95% CI 37-70).

14.3 Process Evaluation - Paper III

We conducted 44 individual interviews. Twelve of these were with children, six with the school heads, ten with grade five teachers, thirteen with parents, three district education officers (two individual interviews, one joint interview), and five focus group discussions (three with grade five children and two with teachers). We also observed at least two of the lessons taught in each of the six schools selected for the process evaluation.

Our findings suggest that the IHC resources were considered important by all of these stakeholders for their ability to address both social and academic issues. Although some of the teachers started out lacking confidence to teach the content, all of them emphasized that the IHC content was important and should be part of the primary school curriculum for science.

The children that attended the lessons found them enjoyable, were motivated and liked both the design and the content of the book, especially the pictures, characters, and games.

The teachers noted that the Teachers' Guide for the IHC lessons (112) was sufficiently flexible to allow for them to employ different ways of doing things in order to accommodate their different teaching styles.

Although the teachers found the IHC lessons to be important, a few still felt that they were an add-on to the already packed primary school curriculum.

The teachers who participated in the trial were motivated by the relevance of the content and strived to create a positive learning environment for the children.

Time was a major constraint for many teachers as a result of other competing priorities. While all of them were able to complete all nine lessons, many felt this was not always to their satisfaction.

Another major constraint was children's absenteeism. This resulted mainly from parents' failure to pay the school fees on time, inability to find appropriate means to travel to school during the rainy season, extra responsibilities for household chores, and poor health or illness.

Teachers felt that in order for the IHC lessons to be successfully scaled up, it is necessary to involve parents, all teachers, and the community members. In addition, it was suggested that it was important to collaborate with officials at the Ministry of Education and the National Curriculum Development Centre to ensure the IHC lessons are incorporated into the primary school curriculum.

Many of the participants we interviewed (teachers, parents, and the children) supported spreading the IHC programme to other schools and to include both younger and older age groups.

The participants identified several benefits of the intervention. Both teachers and children confirmed that they learned important lessons that improved their decision making. Some of the teachers and children also noted improvements in English and numeracy.

Although the majority of teachers enjoyed teaching the IHC lessons, a few reported having experienced stress from teaching something new and it being an add-on to their usual subjects.

14.4 One year post intervention (Paper IV)

A total of 6,787 children completed the one year follow-up test. More children (62%) in the intervention group completed the test than (45%) in the control schools. The proportion of girls (55%) and the median age of children (12,25th to 75th percentile: 10 to 14) in the two groups was similar.

Only 53% of schools in the trial provided data on the secondary outcome of school attendance, 76% of the schools provided data on examination scores for the intervention term, and 83% of the schools provided data for the term after the intervention.

Primary outcomes

The average score for children in the intervention schools was 68.7% compared to 53.0% in the control schools. The adjusted mean difference (based on the regression analysis) was 16.7% (95% CI 13.9-19.5; p<0.00001).

In the intervention group, 80.1% of the children had a passing score (≥ 13 out of 24 correct answers), compared to 51.5% in the control group. The adjusted difference (based on the odds ratio from the logistic regression analysis) was 39.5% more children who passed (95% CI 29.9-47.5) in the intervention than in the control group.

Secondary outcomes

- In the intervention group, 28.9% of the children had a score indicating mastery of the 12 key concepts (≥ 20 out of 24 correct answers) compared to 4.9% of the children in the control schools. The adjusted difference was 25.0% more children in the intervention schools who mastered the concepts (95% CI 23.2- 26.5).
- For each concept the proportion of children who correctly answered both questions for each concept was higher in the intervention schools than in the control schools, including for the concept not included in the primary school resources (p<0.0001 for all 13 concepts after a Bonferroni correction for multiple comparisons).
- Children in the intervention schools were more likely to respond that they would find out what the basis for a claim was compared to children in the control schools (adjusted difference 8.1%, 95% CI 3.7-12.6) and to participate in a research study if asked (adjusted difference 7.7%, 95% CI 2.0- 13.5). This is consistent with what we found immediately after the intervention.
- There was little difference between the two groups on how likely the children were to find out if a claim was based on research (adjusted difference 2.6%, 95% CI -1.9-7.2). This is in contrast with what we found immediately after the intervention (10.8%, 95% CI 6.3-15).
- Children in the intervention schools were more likely to find it easy to assess whether a claim is based on research than children in the control schools (adjusted difference 14.8%, 95% CI 8.9-20.5). This is consistent with what we found immediately after the intervention.
- Children in the intervention were also more likely to consider it easy to find information about treatments based on research (adjusted difference 7.2%, 95% CI 2.6-11.5), whereas a year earlier we had detected little if any difference.

- Intervention children were more likely to report hearing one or more treatment claims daily or weekly compared to children in the control schools (adjusted difference 7.0%, 95% CI 0.5- 12.9).
- In the intervention arm, 98.7% of the teachers had a passing score compared to 85.9% in the control group (adjusted difference 8.6%, 95% CI 1-55.5). While 67.9% teachers in the intervention arm achieved a mastery score ($\geq$ 20 out of 24 correct answers) compared to 21.9% in the control group (adjusted difference 46.3%, 95% CI 31.5-56.6). These results are similar to what we found at the end of the immediately after the intervention.

15 DISCUSSION

15.1 SUMMARY OF THE KEY FINDINGS

It is possible to teach critical thinking skills in primary schools in a low-income country

The trial reported in this thesis, which evaluated the Informed Health Choices (IHC) primary school resources in Uganda, demonstrated that children as young as 10 years old can be taught principles of reasoning, based on empirical evidence, logic, and scientific reasoning. In this study, 69% of the children in the intervention arm achieved a passing score compared to only 27% of the children in the control arm (16).

Although the intervention lasted only one school term, the large effect that the intervention had was sustained after one year. In fact, the proportion of children with a passing score in the intervention arm increased from 69% in the evaluation done immediately after the intervention to 80% after one year(16).

With rapid technological transformations and information overload, it is increasingly important that children are able to learn to think critically, and not just to repeat a list of facts from memorization. Children must be critical thinkers who can make sense of information, analyse, compare, contrast, make inferences and demonstrate higher order thinking skills in order to successfully navigate the information age.

Teaching children critical thinking skills should no longer be considered a “nice to have” part of their formal and informal education. It is a “must have” in Uganda, as well as in education systems worldwide. As such, recognised standards and methods of teaching critical thinking should be established in schools, starting from primary school.

In a systematic review that analysed 117 studies of the effects of strategies for teaching critical thinking, Abrami and colleagues found that explicit instruction was the teaching approach with the strongest empirical support for improving critical thinking skills (60). Explicit instruction, where children were taught to reason and solve problems, was more effective than strategies where teachers asked students to solve problems without giving them explicit instruction. The large effect of the IHC primary school intervention, which used explicit instruction, is consistent with those findings.

Recent research clearly indicates that the ability to think critically is essential for attainment of better educational outcomes and as a lifelong skill (113). As such, education systems are required to integrate the teaching of critical thinking skills into their existing curriculums. Low levels of educational attainment impact negatively of on later life chances, especially in relation to participation on the labour market. This has long-term consequences not only for the individual, but also for the country's economic prosperity (114).

Use of a user-centred design approach resulted in materials that were considered relevant, valuable and appropriate by end users

We employed a user-centred design approach with multiple iterations of feedback from teachers and children to inform the development of the intervention. This resulted in learning resources that both teachers and children experienced positively. Key findings of pilot and user-testing were that we needed to simplify and reduce the amount of material included in the earlier prototypes of the learning resources. Although we identified 24 key concepts that could and should be taught to primary school children, this was too much to teach in a single school term. Consequently, the learning resources that we evaluated included only 12 key concepts and they began with very simple and basic explanations. This included explanations of, for example, what "health" is. This helped to ensure that the children understood the lessons and that they had a foundation that they could build on from chapter to chapter, and in the future.

These findings are consistent with the logic underlying a "spiral curriculum". Bruner defined four main characteristics of a spiral curriculum (115):

1. Existence of increasing levels of difficulty and or depth throughout the curriculum
2. Re-visitation of topics throughout the curriculum with increased complexity
3. New learning should be related to previous learning
4. Gradual increase in learners' competence until the overall learning objectives are achieved.

The effectiveness of the intervention may also be attributable in part to re-visitation of what was learned in the previous lesson at the beginning of each subsequent lesson, and again in the last lesson, which reviewed all of the key concepts that were included. Each lesson also included opportunities to apply what was learned in individual exercises and class activities. The process evaluation indicates that the children also applied what they were learning outside of the classroom. Spaced practice, with intervals between learning sessions, has been found to improve long-term retention (116, 117). This may have contributed to the high level of retention, documented in the one year follow-up study.

Teaching additional key concepts, long-term retention, and changes in behaviour likely require a spiral curriculum that extends over several years. This, in turn, may require changes in the national curriculum to ensure that there is adequate time for this and that it is prioritised. This is supported by our process evaluation. The main barrier to scaling up use of the IHC resources that participants in that study identified was the need to incorporate the lessons into the national curriculum.

Scaling up and sustaining teaching critical thinking in primary schools in low-income countries requires an affordable intervention

The majority of parents, teachers and children who participated in the process evaluation emphasised that the IHC primary school resources were relevant. They also believed that efforts to scale up use in all primary schools and classes should be intensified. Although they identified the need to change the national curriculum as the most important barrier to doing this, a bigger barrier to scaling up use of the learning resources that we evaluated is their cost.

We estimated that the cost of the IHC primary school intervention was approximately 4 US dollars per child. In a low-income country such as Uganda, where the annual government expenditure per primary school student is approximately 29.4 US dollars, this is not affordable. Over 90% of the cost of the intervention was for printing and distribution of the learning resources (Appendix 1 in Paper II). As computers become more available, the marginal cost of providing the learning resources digitally would be minimal. There are large-scale, government-supported educational tablet initiatives in several countries in Africa and the Middle East (118) which could make scaling up the use of digital learning resources feasible. Unfortunately, there is currently no such program in Uganda. On the other hand, computers are widely available in Ugandan secondary schools. Thus, it might make sense to develop and evaluate learning resources for secondary schools, which could feasibly be scaled up, even though it currently would not be possible for those resources to build on what was learned in primary school.

16 STRENGTHS AND LIMITATIONS

16.1 STRENGTHS

Using a user-centred design approach resulted in an intervention that the end users experience positively

We employed a user-centred design approach that enabled user engagement throughout the development process. Our multidisciplinary team consisted of researchers with diverse backgrounds that included information design, journalism, health service research, and research methodology. We worked closely with the teachers' network, policy makers, education researchers, and children to generate ideas, test prototypes, and gather feedback. This helped to create resources that resonated with the end users.

The IHC primary school resources were developed over a period of three years. This allowed us time to generate and prototype ideas; and to iteratively design, pilot, user test, analyse and redesign those prototypes. In addition, the earlier versions of the resources were piloted and user-tested in Kenya, Rwanda, and Norway. This helped to ensure that they were usable in settings outside of Uganda.

As a result of the iterative revisions of the IHC primary school resources, children and teachers experienced the resources as useful, easy to use, understandable, credible, desirable and well suited to their context. Pilot and user testing in Kenya, Rwanda, and Norway, as well as ongoing efforts in other countries, indicate that the learning resources can be translated and contextualised for use in other settings without major changes.

Conducting a large cluster-randomised trial with one year follow-up provided compelling evidence of the effectiveness of the intervention

We evaluated the effects of the IHC primary school intervention in a large cluster randomised trial, measuring outcomes shortly after the intervention and again after one year. Despite some limitations, this pragmatic trial provided robust evidence of the effectiveness of the intervention.

Using a multi-method process evaluation provided a solid basis for understanding how and why the intervention worked and insights into potential effects and limitations.

We used a multi-method approach for the process evaluation, which included both quantitative and qualitative data. This included routinely collected data on attendance and performance, lesson evaluation forms, observations, face-to-face interviews, and focus groups. Using a multi-method approach increased the robustness of our findings because results were strengthened through triangulation. Cross-validation was achieved through the convergence of different kinds and sources of data which were found to be in harmony (119). In addition, we applied a modified CERQual approach to explicitly assess our confidence in each finding. This indicated high confidence in most of our findings.

16.2 LIMITATIONS

There were limitations to the outcomes we measured.

In our systematic review of existing measurement instruments (2), we found none that would be suitable for use of our trial in Uganda. Consequently, we used a tool that was developed by us. Although we used selected multiple-choice questions from the CLAIM Evaluation Tools Database that had been judged by independent research methodologists to have face validity, and by end users to be relevant and understandable, the test used was aligned with the intervention. It was “treatment inherent”, in that the test measured the ability to apply the concepts that the resources were intended to teach. Treatment-inherent outcome measures are associated with larger effect sizes than independent measures (120). As such, it is difficult to compare our findings to the findings of studies that used independent measures, such as literacy and numeracy tests.

Another limitition is that we were not able to measure the effects of the intervention on actual health choices and outcomes; or to measure long-term effects on decision-making, health behaviours or health. It is unlikely that a one-off intervention like ours has large long term effects on behaviours or health outcomes. However, as noted above, we believe it lays the foundation for subsequent learning and the development of dispositions (frequent and voluntary habits of thinking and doing) for thinking critically about treatment claims and choices.

A third limitation of the primary outcome measure was the use of multiple-choice questions to measure critical thinking skills. From a researcher’s (and a teacher’s) point of view, using multiple-choice questions is advantageous with respect to ease of scoring, fast return of scores in large classrooms, the potential to ask more questions than with essays, and objectivity (less judgement required to score answers). However the use of multiple-choice questions has limitations for assessing critical thinking skills(121, 122). Constructed (subjective) response questions enable students to express what they know at different thinking levels, but grading such questions is more subjective (requires more judgement), there is more intergrader variability, and more time is required for grading (123).

Students often view multiple-choice tests as easier than essay exams (122). They require less time and effort to prepare for, and the availability of options to choose from is comforting (124). Students tend to expect more high-level questions from constructed response exams and employ more deep-learning strategies in preparation for them (125). Thus, the ability to draw conclusions about acquisition of higher order thinking skills based on our primary outcome measure is limited. However, it is unlikely that preparation for the test that we used biased the results in favour of the intervention. Neither the teachers nor the children were exposed to the outcome measure or similar multiple-choice questions until they took the test at the end of the school term when the intervention was delivered.

There are major barriers to scaling up the intervention in Uganda

The IHC primary school resources were found to be effective, teachers and parents considered the content to be very important, and children greatly enjoyed and valued the lessons. School authorities, teachers, and parents all felt that the intervention should be scaled up. However, the cost of the intervention and the need to find time for it in an already over-crowded curriculum were found to be major barriers to scaling up the intervention. It is possible that if we had given this more consideration in the development phase, we might have designed an intervention that would cost less and be easier to integrate in the current curriculum. On the other hand, if we had done that, although it might have been more pragmatic to scale up use of the intervention, the intervention might have been less effective.

17 ETHICAL CONSIDERATIONS

Obtaining assent and parental consent is the first step in involving children as participants in research. Children as young as seven years old have demostrated an ability to understand the essential elements of research (126, 127). Thus it may be possible and important to obtain their assent to participate in research. Variability in understanding reported by some reseachers can be attributed to the complexity of the language used in the assent forms, and may not reflect developmental differences (127). The challenge for researchers is to find innovative ways in which to communicate information in a manner that children can understand (128).

In our study, the majority of children that participated did not have English as their first language and many had poor literacy skills. Whilst the parental consent forms were written in two languages (English and Luganda), the children's assent forms, although less detailed, were written in English only. This was based on the assumption that grade five children had some understanding and an adequate command of the English language.

Our intervention posed minimal risk and no more risk than other teaching materials, most of which have not been evaluated. Only children who participated in individual interviews and focus group discussions during user testing and piloting were required to obtain parental consent before offering their assent to participate. We did not obtain assent from individual children or consent from their parents for participation in the trial. Headteachers decided whether their school was to participate and consented on behalf of the school, and the teachers consented to participante. We informed children and parents about the purpose of the trial, but we did not seek their assent or consent. This is not different from nearly everything else that is done in primary schools, where individual students and their parents do not assent or consent to what is taught, how it is taught, or how academic achievement is measured.

Assent processes provide an opportunity to assess children's understanding of the research and for children to interact with the researchers and build trust. This is especially important when dealing with older children and adolescents and for medical research. Older children and adolescents may desire and expect privacy and autonomy from their parents or legal guardians (129). Studies of consent and assent processes have been done in populations of ill children (126), but little research has been done involving children in other contexts, particularly in education settings in low-income countries. In this research, we have shown that it is possible to explain key concepts underlying evaluations of the effects of treatments to young children in settings with few resources. It is possible that the approach that we have used and some of the resources that we have developed could be adapted and used to ensure that children, young people, and adults with little formal education understand the research when they are invited to participate.

Prior consultations with representatives on the teachers' network during the user testing and piloting phase provided guidance for the selection of consent and assent procedures within this context. Although, this type of consultation can be time consuming and burdensome for researchers, it demostrated the study team's commitment to respecting local norms.

We engaged children, parents, teachers, and school authorities in the development of the intervention over a three year periond. We also engaged with them prior to and during the trial. We visited each school and organised workshops for the teachers at the start of the trial to explain plans for the trial to them. This resulted in mutual respect between the research team and the study participants, and cultivated trust. In the trial results, an overwhelming number of participants, including both children and adults in the intervention group, confirmed a willingness to participate in research in the future. In addition, children that participated in the individual interviews and focussed group discussions reported positive views about their research experiences. However, children generally tend to positively respond when subjectively asked about their experiences (130). Further research is needed to investigate children's views on research participation in this context.

18 IMPLICATIONS

18.1 IMPLICATIONS FOR PRIMARY EDUCATION IN UGANDA

Critical thinking is currently being taught to a limited extent in Uganda. Our trial shows that it is possible to do this and findings of the process evaluation show that school authorities, teachers, parents, and children agree that critical thinking should be taught to a greater extent than it currently is. Health is an ideal topic for teaching critical thinking because it is important and relevant to everyone, including children. However, before use of the IHC primary school resources can be scaled up in Uganda, two things must happen.

First, the national curriculum must be changed so that teaching critical thinking is made a core component instead of a mere an add-on. Although it was clearly indicated by the various stakeholders (head teachers, parents), end-users (teachers) and educational authorities (district educational officers and curriculum content experts at the National Curriculum Development Centre) during our consultations that the IHC resources were highly relevant to primary school children, there were varying issues of concern. Curriculum experts were initially concerned about the appropriateness of the content and its' compatibility with the curriculum. District education officers were concerned about time required by the teachers to prepare and how it would affect teaching of the core subjects. Teachers were worried about the extra load this would add onto their already busy schedules. Based on our experience, such concerns have to be identified early on and addressed with solutions clearly outlined for interventions for teaching critical thinking in schools to be effective.

Second, either the IHC learning resources must be made affordable or alternative learning resources must be developed and evaluated. The most likely way in which the IHC learning resources could become affordable is through digitalisation, which presupposes that computers become available in primary schools. Computers are increasingly available in primary schools in Kenya and Rwanda, and there is a nascent program for the same in Uganda. However, the government currently is not supporting such a program in Uganda.

As described in Paper I, we generated a large number of ideas, working together with the teachers' network, and prototyped several, none of which appeared viable. A key factor that led us to developing the textbook and teachers' guide was that teachers needed this support to enable them to teach something that was new to them. Alternative resources that did not provide this support are unlikely to be effective. Consequently, at this time, the strategy that seems most promising is to develop and evaluate digital learning resources for secondary schools, where computers are widely available. We will attempt this in another research project funded by the Research Council of Norway.

18.2 IMPLICATIONS FOR OTHER COUNTRIES

There has been widespread interest in translating, contextualising, and testing the IHC primary school resources in other countries. Pilot testing of translated versions in Kenya and Rwanda indicate that only minor changes are needed for them to be used in those countries. Pilot testing of the English version in Ireland also found that only minor chages are needed there. This is similar to what we found piloting an earlier version of the resources at an international school in Norway. At the school in Norway, they integrated the IHC lessons into their science curriculum. They provide the *Health Choices Book* to students on tablets, using the PDF version that is freely available on the IHC website. At both the international school in Norway and in Ireland the teachers were sceptical about the book to begin with.They thought that the children would not relate to it culturally and that it might be too simple for the children. In both cases, the children enjoyed the book and did not have a problem relating to the story in it.

Based on this experience, we advise people to first pilot and user test the IHC primary school resources before deciding whether major changes are needed. We have developed a guide for translating and contextualising the resources for use in other countries, which includes this advice (131).

18.3 IMPLICATIONS FOR RESEARCH

The four papers included in this thesis provide strong support for the value of employing a user-centred design process to develop learning resources. We will continue to use that approach ourselves and would encourage others to use this approach to develop learning resources.

Priorities for future research into how best to enable students to think critically about treatment claims and choices include:

- Developing a spiral curriculum that maps out which key concepts to begin with and at what age; the order in which new key concepts should be taught, building on previous lessons; the intervals between lessons; and the goals, including competences and dispositions (132).
- Undertaking "market analyses" to inform adaptation of the IHC primary school resources and the development of new learning resources. This includes identifying where teaching critical thinking about health best fits in the curriculum and clarifying conditions for introducing this into schools, including the availability of time, who the decision-makers are, and what influences their decisions (133).
- Developing and evaluating new learning resources for each interval of a spiral curriculum. Ideally, these resources should either be tested at scale, or the resources that are developed should be designed so that their use can easily be scaled up, if they are found to be effective.
- Developing and validating outcome measures for relevant dispositions and decision-making behaviours; measuring effects on other outcomes, including academic achievement and application of the IHC Key Concepts to other types of interventions (besides health interventions) and decisions; and measuring longer term effects (beyond one year).

Results in relation to other research

We have not found any directly comparable studies of interventions such as this in primary schools with a positive effect on critical thinking in low-income countries. The effect size for the evaluation study (a standardised mean difference of 1.16; 95% CI 1.00 to 1·32) is well above the average effect size reported for other critical thinking interventions (0.33; 95% CI 0.31 to 0.34) (60). It is also above the average effect size found for (treatment inherent) outcome measures developed by the study's authors (0·65; 95% CI 0·52 to 0·78) (60). It is also larger than any of the effect sizes reported in a systematic review of interventions to improve learning in primary schools in low-income and middle-income countries for both interventions with teacher training and instructional materials (134) and for structured pedagogy programmes (135). However, most of those studies measured scores on standard reading or math tests, not on critical thinking.

In a systematic review of educational interventions for teaching critical thinking (59), many of the included studies measured outcomes immediately following the intervention. Only a few of the studies measured outcomes a short while (2 to 6 weeks) after the intervention. In our follow-up study, we found that children retained what was learned for at least one year. Longer term effects and effects on attitudes and actual behaviours remain uncertain.

As far as we are aware, there are currently no studies measuring students' critical appraisal behaviours outside the classroom environment, which is the ultimate goal for teaching critical thinking skills. Evaluations of educational interventions in schools have typically measured short-term cognitive learning outcomes (37). Although we measured self-reported behaviours in our follow-up study and explored the potential outcomes in the process evaluation, more work is neeed. Future research should aim to measure longer-term effects on dispositions and behaviours, as well as on critical thinking skills.

In the systematic review of educational interventions to teach critical thinking (60), Abrami and colleagues found that active learning strategies promoted critical thinking in young people. Our findings are consistent with that finding. However, the approaches used to teach critical thinking in many of those studies differed from the traditional authoritative approach that is familiar to many teachers in Uganda and other low-income countries. Nonetheless, we found in the process evaluation that the teachers felt that the approach used in our intervention was appropriate for them, and this may have contributed to the effectiveness of our intervetnion.

Many teachers may be uncomfortable teaching critical appraisal skills because they have not learned those skills themselves (37). This is consistent with the findings of our pilot and user testing and our process evaluation. Another factor that may have contributed to the effectiveness of our intervention was the scaffolding that we provided for both the teachers and the students through the teachers' guide and the children's textbook. This enabled the teachers to teach these skills to the children for the first time, despite their initial discomfort. The large effect of the intervention on the teachers' mastery of the key concepts after teaching them for one school term supports this conclusion. In our process evaluation study, the majority of the teachers stated that the IHC content was new to them, with some expressing concern about their understanding of the content and their ability to teach it.

Despite these initial concerns, 94% of the teachers agreed or strongly agreed in their end of term assessments that they understood the cintent of the lessons, and 97% of them responded that they learned a great deal.

19 CONCLUSION

Our intervention was effective in teaching primary school children in a low-income country to appraise health information. This research documents the potential to teach critical thinking skills to children in low-income settings, despite large student-teacher ratios and low literacy levels. It also documents the value of a user-centred approach to developing educational interventions. However, constrained educational budgets limit the possibility of scaling up the intervention that we evaluated. Further research is needed to develop and evaluate interventions implemented at scale. This research provides a solid basis for doing that.

In addition, it will be necessary to incorporate teaching critical thinking, and critical appraisal of treatment claims specifically, into the national curriculum. It also is important to do this as a spiral curriculum over several years, rather than as a one-off intervention. Again, this research provides a solid basis for that work and for undertaking rigorous long-term evaluations of such an approach.

In summary, this research is an important initial step towards developing, evaluating, scaling up, and sustaining the use of interventions that will equip young people with the skills they need to make informed health choices as future patients, health professionals, policy makers and citizens.

20 REFERENCES

1. Austvoll-Dahlgren A, Oxman AD, Chalmers I, Nsangi A, Glenton C, Lewin S, et al. Key concepts that people need to understand to assess claims about treatment effects. Journal of Evidence Based Medicine. 2015;8:112-25.
2. Austvoll-Dahlgren A, Nsangi A, Semakula D. Interventions and assessment tools addressing key concepts people need to know to appraise claims about treatment effects: a systematic mapping review. Systematic Reviews. 2016;5(1):215.
3. Parker R, Baker D, Williams M, Nurss J. The test of functional health literacy in adults. J Gen Intern Med. 1995;21:651-7.
4. Davis TC, Crouch MA, Long SW, Jackson RH, Bates P, George RB, et al. Rapid assessment of literacy levels of adult primary care patients. Fam Med. 1991;23:433-5.
5. Ishikawa H, Nomura K, Sato M, Yano E. Developing a measure of communicative and critical health literacy: a pilot study of Japanese office workers. . Health Promot Int. 2008;23(269-74).
6. Osborne R, Batterham R, Elsworth G, Hawkins M, Buchbinder R. The grounded psychometric development and initial validation of the Health Literacy Questionnaire (HLQ). . BMC Public Health. 2013;13(658).
7. Sorensen K, Van den Broucke S, Pelikan J, Fullam J, Doyle G, Slonska Z, et al. Measuring health literacy in populations: illuminating the design and development process of the European Health Literacy Survey Questionnaire (HLS-EU-Q). BMC Public Health 2013;13(948).
8. Giacomin J. What is human centred design? Design J 2014;17:606-23.
9. Austvoll-Dahlgren A, Guttersrud O, Semakula D, Nsangi A, Oxman AD, Chalmers I, et al. The development of the "Claim Evaluation Tools": assessing critical thinking about effects. BMJ Open. 2017;7(5:e013185).
10. Austvoll-Dahlgren A, Guttersrud Ø, Semakula D, Nsangi A, Oxman AD, al. E. Measuring ability to assess claims about treatment effects: A latent trait analysis of the CLAIM Evaluation Tools using Rasch modelling. BMJ Open. 2017;0:e013185. .
11. Semakula D, Nsangi A, Øystein G, Oxman AD, Sewankambo NK, Austvoll-Dahlgren A. Measuring ability to assess claims about treatment effects in English and Luganda: evaluation of multiple-choice questions from the “Claim Evaluation Tools” database using Rasch modelling. Informed Health Choices Working Paper, 2017 ISBN 978-82-8082-830. 2017;https://www.informedhealthchoices.org/wp-content/uploads/2016/08/Claim-2nd-Rasch-analysis-in-Uganda-2017-03-17.pdf.
12. Nsangi A, Semakula D, Oxman AD, Austvoll-Dahlgren A, Oxman M, Rosenbaum S, et al. Effects of the Informed Health Choices primary school intervention on the ability of children in Uganda to assess the reliability of claims about treatment effects: a cluster-randomised controlled trial. Lancet. 2017;390:374-88.
13. Semakula D, Nsangi A, Oxman AD, Oxman M, Austvoll-Dahlgren A, Rosenbaum SE, et al. Effects of the Informed Health Choices podcast on the ability of parents of primary school children in Uganda to assess claims about treatment effects: a randomised controlled trial Lancet. 2017;390(10092):389-98.
14. Davies A, Gerrity M, Nordheim L, Peter O, Opiyo N, Sharples J, et al. Measuring ability to assess claims about treatment effects: establishment of a standard for passing and mastery. . IHC working paper, http://wwwinformedhealthchoicesorg/wp-content/uploads/2016/08/Claim-cut-off-IHC-Working-Paper-2017-01-09pdf 2017;Accessed 30 October 2018.
15. Semakula D, Nsangi A, Oxman AD, Oxman M, Austvoll-Dahlgren A, Rosenbaum S, et al. Effects of the Informed Health Choices podcast on the ability of parents of primary school children in Uganda to assess the trustworthiness of claims about treatment effects: one-year follow up of a randomised trial. Trials. 2020;21(1):187.
16. Nsangi A, Semakula D, Oxman AD, Austvoll-Dahlgren A, Oxman M, Rosenbaum S, et al. Effects of the Informed Health Choices primary school intervention on the ability of children in

Uganda to assess the reliability of claims about treatment effects, 1-year follow-up: a cluster-randomised trial. Trials. 2020;21(1):27.
17. Nsangi A, Semakula D, Glenton C, Lewin S, Oxman AD, Oxman M, et al. Informed health choices intervention to teach primary school children in low-income countries to assess claims about treatment effects: process evaluation. BMJ Open. 2019;9(9):e030787.
18. Semakula D, Nsangi A, Oxman A, Glenton C, Lewin S, Rosenbaum S, et al. Informed Health Choices media intervention for improving people’s ability to critically appraise the trustworthiness of claims about treatment effects: a mixed-methods process evaluation of a randomised trial in Uganda. BMJ Open. 2019;9(12):e031510.
19. Glenton C, Paulsen EJ, Oxman AD. Portals to Wonderland: health portals lead to confusing information about the effects of health care. BMC Med Inform Decis Mak. 2005;5(7).
20. Cooper B, Lee W, Goldacre B, Sanders T. The quality of dietary advice given in UK national newspapers. Public Underst Sci. 2011;21:664-73.
21. Woloshin S, Schwartz LM, Moncur M, Gabriel S, Tosteson AN. Assessing values for health: numeracy matters. . Medical Decision Making. 2001;21(5):382-90.
22. Sillience E, Briggs P, Harris PR, Fishwick L. How do patients evaluate and make use of online health information? Soc Sci Med. 2007;64(9):1853-62.
23. Schwartz L, Woloshin S, Black W, Welch H. The role of numeracy in understanding the benefit of screening mammography. Ann Intern Med. 1997;127(11):966-72.
24. Lloyd A. The extent of patients' understanding of the risks of treatments. Quality in Health Care. 2001;10(1):114-8.
25. Glenton C, Nilsen ES, Carlsen B. Lay perceptions of evidence-based information- a qualitative evaluation of a website for back pain sufferers. BMC Health Services Research. 2006;6(34).
26. Oxman AD, Austvoll-Dahlgren A, Garratt A, Rosenbaum S. Understanding of key concepts relevant to assessing claims about treatment effects: a survey of Norwegian adults. IHC Working Paper. 2017.
27. Eysenbach G, Kohler C. How do consumers search for and appraise health information on the World Wide Web? A qualitative study using focus groups, usability tests and in-depth interviews. BMJ Open. 2002;324:373-7.
28. Verhoef MU, Mulkins A, Carlson LE, Hilsden RJ, Kania A. Assessing the role of evidence in patients' evaluation of complimentary therapies: a qualitative study. Integr Cancer Ther. 2007;6:345-53.
29. Academy of Medical Sciences. Enhancing the use of scientific evidence to judge the potential benefits and harms of medicines. London: Academy of Medical Sciences,. 2017.
30. Wooten D, ReedII A. Informational Influence and the Ambiguity of Product Experience: Order Effects on the Weighting of Evidence. Journal of Consumer Psychology. 1998;7(1):79-99.
31. Aronson E, Wilson TD, Akert AM. Social Psychology. 5th Edition, Upper Saddle River, NJ: Prentice Hall. 2005.
32. Cialdini BR, Wosinska W, Barett DW, Gornik-Durose M. "Compliance with a request in two cultures: The differential influence of social proof and commitment/consistency on collectivists and individualists". Personality and Social Psychology Bulletin. 1999;25(10):1242-53.
33. Chalmers I, Glasziou P, Badenoch D, Atkinson P, Austvoll-Dahlgren A, A O. Evidence live 2016: promoting informed healthcare choices by helping people assess treatment claims. BMJ http//blogsbmj/com/bmj/2016/05/26/evidence-live-2016- promoting-informed-healthcare-choices-by-helping-people-assess-treatment-claims. 2016.
34. Woloshin S, Schwartz LM. Media reporting on research presented at scientific meetings: more caution needed. Medical Journal of Australia. 2006;184(11):576-80. Epub Published online: 5 June 2006.
35. Evans I, Thornton H, Chalmers I, Glasziou P. Testing Treatments. 2nd edition ed: Pinter and Martin; 2011.

36. Chalmers I, Oxman AD, Austvoll-Dahlgren, Ryan-Vig S, Pannell S, Sewankambo N, et al. Key Concepts for Informed Health Choices: a framework for helping people learn how to assess treatment claims and make informed choices. Evidence Based Medicine. 2018;23:29-33.
37. Nordheim L, Gundersen MW, Espehaug B, Guttersrud O, Flottorp S. Effects of school based educational interventions for enhancing adolescents' abilities in critical appraisal of health claims. a systematic review. Plos One. 2016;11(0161485).
38. Ybarra M, Suman M. Reasons, assessments and actions taken: sex and age differences in uses of Internet health information. Health Educ Res. 2008;23:512-21.
39. Shekelle PG, Pronovost PJ, Wachter RM, et al. The top patient safety strategies that can be encouraged for adoption now. Annals of Internal Medicine. 2013;158(5_Part_2):365-8.
40. Steinman MA. Using patients to promote evidence-based prescribing: comment on "Communicating uncertainties about prescription drugs to the public". Archives of internal medicine. 2011;171(16):1468-9.
41. Fønhus MS, Dalsbø TK, Johansen M, Fretheim A, Skirbekk H, Flottorp SA. Patient-mediated interventions to improve professional practice. Cochrane Database of Systematic Reviews 2018(9).
42. Ennis RH. Critical thinking: reflection and perspective: part I,. Inquiry 2011. 2011;26:4-18.
43. Organization WH. Health and Millenium Development Goals; 1.World health, 2. Health priorities, 3. Delivery of health care, 4. Cost of illness, 6. Goals, 7. Social justice. WHO Library Cataloguing-in-Publication Data. 2005.
44. Austvoll-Dahlgren A, Barends E, Boruch R, Brennan M, Chalmers I ea. Key concepts for informed choices. Nature (In Press). 2019.
45. Sharples JM, Oxman AD, Mahtani KR, Chalmers I, Oliver S, Collins K, et al. Critical thinking in healthcare and education. BMJ. 2017;357.
46. Smith VG, Szymanski A. Critical Thinking: More than Test Scores. International Journal of Educational Leadership Preparation. 2013;8:16-25.
47. Bloom B, Englehart M, Furst E, Hill W, Krathwohl D. Taxonomy of educational objectives: The classification of educational goals. . Handbook I: Cognitive domain New York, Toronto: Longmans, Green 1956.
48. National Population Council. State of Uganda Population Report 2017; "Transforming Uganda's Economy: Opportunities to Harness the Demographic Dividend for Sustainable Development''. 2017;17.
49. Internet World Stats. Internet World Stats 2017. Internet World StatsOrg. 2017.
50. Park E, Kwon M. Health-Related Internet Use by Children and Adolescents: Systematic Review. Journal of medical Internet research. 2018;20(4):e120-e.
51. Chen L, Nath R. Understanding the underlying factors of Internet addiction across cultures: A comparison study. Electronic Commerce Research and Applications. 2016;17:38-48.
52. Oxman AD, Paulsen EJ. Who can you trust? A review of free online sources of "trustworthy" information about treatment effects for patients and the public. BMC Medical Informatics and Decision Making. 2019;19(1):35-.
53. Sørensen K, Van den Broucke S, Fullam J, Doyle G, Pelikan J, Slonska Z, et al. Health literacy and public health: A systematic review and integration of definitions and models. BMC Public Health. 2012;12(1):80.
54. Stead M, Eadie D, Gordon D, Angus K. Hello, hello- it's English I speak"! a qualitative exploration of patients' understanding of the science of clinical trials. Journal of Medical Ethics. 2005;31(11):664-9.
55. Sheridan SL, Halpern DJ, Viera AJ, Berkman ND, Donahue KE, Crotty K. Interventions for individuals with low health literacy: a systematic review. Journal of Health Communication. 2011;16(Suppl 3):30-54.
56. Berkman ND, DeWalt DA, Pignone MP, Sheridan SL, Lohr KN, Lux L. Literacy and Health Outcomes. Summary Evidence Report/Technology Assessment. 2004(87):8.

57. Betz C. Health Literacy: The missing link in the provision of health care for children and their families. Journal of Pediatric Nursing. 2007;22(4):257-9.
58. Nutbeam D, McGill B, Premkumar P. Improving health literacy in community populations: a review of progress. Health Promot Int 2018;33:901-11.
59. Cusack L, Del Mar CB, Chalmers I, Gibson E, Hoffmann TC. Education interventions to improve people's understanding of key concepts in assessing the effects of health interventions: a systematic review. Systematic Reviews. 2018;7(68).
60. Abrami PC, Bernard RM, Borokhovski E, Waddinton DI, Wade CA, Persson T. Strategies for teaching students to think critically; a meta analysis. Rev Edu Res. 2015;85:275-314.
61. Snilstveit B, Stevenson J, Phillips D, et al. Interventions for improving learning outcomes and access to education in low- and middle-income countries: a systematic review. . London: International Initiative for Impact Evaluation. 2015.
62. Evans DK, Popova A. What really works to improve learning in developing countries? An analysis of divergent findings in systematic reviews. . Policy Research Working Paper Washington DC World Bank. 2015;7203.
63. Austvoll-Dahlgren A, Nsangi A, Semakula D. Interventions and assessment tools addressing key concepts people need to know to appraise claims about treatment effects: a systematic mapping review. Systematic Reviews. 2016;5(215).
64. National Health and Medical Research Council. How to Use the Evidence: Assessment and Application of Scientific Evidence. Canberra 2000. 2000.
65. Katrak P, Bialocerkowski AE, Massy-Westropp N, Kumar VS, Grimmer KA. A systematic review of the content of critical appraisal tools. BMC Medical Research Methodology. 2004;4(1):22.
66. Cusack L, Del Mar CB, Chalmers I, Hoffmann TC. Educational interventions to improve people's understanding of key concepts in assessing the effects of health interventions: a systematic review protocol. Systematic Reviews. 2016;5(37).
67. Duschl RA, Schweingruber A, Shouse AW, (editors). Committee on Science Learning, Kindergarten Through Eighth Grade. How children learn science. Taking science to school: learning and teaching science in grades K-8. The National Academies Press, Washington, DC: . 2007.
68. VanTassel-Baska J, Bracken B, Feng A, Brown E. A Longitudinal Study of Enhancing Critical Thinking and Reading Comprehension in Title I Classrooms. Journal for the Education of the Gifted. 2009;33(1):7-37.
69. Tsai P, Chen S, Chang H, Chang W. Effects of prompting critical reading of science news on seventh graders' cognitive achievement. International Journal of Environmental & Science. 2013;8(1):85-107.
70. McCollister K, Sayler M. Lift the ceiling: Increase rigor with critical thinking skills. . Gifted Child Today. 2010;33(1):41-7.
71. Sandoval WA, Sodian B, Koerber S, J W. Developing children's early compe-tencies to engage with science. Educ Psychol 2014; 49: 139-52. Educ Psychol 2014; 49: 139-52. 2014.
72. Ruddock G, Sainsbury M. Comparison of the core primary curriculum in England to those of other high performing countries. Research Report DCSF-RW048. London: Department for Children, Schools and Families, 2008. Research Report DCSF-RW048 London: Department for Children, Schools and Families, 2008. 2008.
73. Tamusuza A. Leaving School Early, The Quest for Universal Primary Education in Uganda. Journal of statisque african numero 13; November 2011. 2011.
74. Burke K, Beegle K. Why Children Aren't Attending Schools, The Case of Northwestern Tanzania. Journal of African Economies. 2004;13(2):333-55.
75. Deininger K. Does the Cost of Schooling Affect Enrollment by the Poor? Universal Primary Education in Uganda. Economics of Education Review Journal. 2003;22(3):291-305.
76. United Nations. Transforming our world- The Agenda for Sustainable Development Goals. United Nations. 2017.

77. Berkman N, Sheridan S, Donahue K, Halpern D, Crotty K. Low Health Literacy and Health Outcomes: An updated systematic review. Annals of Internal Medicine. 2011;155(2):97-U89.
78. Eichler K, Wieser S, Brügger U. The costs of health literacy: a systematic review. International Journal of Public Health. 2009;54:313-24.
79. Cleland JG, van Ginneken JK. Maternal education and child survival in developing countries: The search for pathways of influence. Social Science & Medicine. 1988;27(12):1357-68.
80. Vikram K, Vanneman R, Desai S. Linkages between maternal education and childhood immunization in India. Social Science & Medicine. 2012;75(2):331-9.
81. Greenaway ES, Leon J, Baker DP. Understanding the association between maternal education and use of health services in Ghana: exploring the role of health knowledge. J Biosoc Sci. 2012;44(6):733-47. Epub 2012/03/01.
82. Mensch BS, Chuang EK, Melnikas AJ, Psaki SR. Evidence for causal links between education and maternal and child health: systematic review. Trop Med Int Health. 2019;24(5):504-22.
83. Klugman J. Human development report 2010: the real wealth of nations: pathways to human development. Palgrave Macmillan, New York. 2010.
84. World Health Organization (2006). Global health observatory data repository. World Health Organization, Geneva Accessed September 7 2018.
85. World Health Organization. World Health Organisation (2006) Global health observatory data repository. World Health Organization, Geneva. 2006.
86. Campbell J, Dussault G, Buchan J, Pozo-Martin F, Guerra Arias M, Leone C, et al. A universal truth: no health without a workforce. World Health Organization; Geneva. 2013.
87. National Curriculum Development Centre 2006a. The National Primary School Curriculum for Uganda Primary 1. The National Curriculum Development Centre (NCDC) Kampala. 2006.
88. Hulya Kosar Altinyelken. Curriculum change in Uganda: Teacher perspectives on the new thematic curriculum. International Journal of Educational Development 2010;30(2010):151-61.
89. Ministry of Education and Sports 2016. Education Statistic Abstract 2016 Ministry of Education and Sports. Ministry of Education and Sports Kampala UgandaMin of Educ and Sports primary_ educn_statistics_abs2016htm (accessed 19 September 2018) http://wwweducationgoug/.
90. Goldstein H. Multilevel Statistical Models. 1995. New York, Halstead Press.
91. Nsangi A, Semakula D, Oxman DA, Sewankambo KN. Teaching children in low income countries to assess claims about treatment effects; A prioritisation of key concepts. Journal of Evidence Based Medicine 2015. 2015;8(4):173-80.
92. Irwig L, Irwig J, Trevena L, Sweet M. Smart Health Choices: Making Sense of Health Advice. London: Hammersmith. 2008.
93. Oxman AD, Guyatt GH, Cook DJ, Jaeschke R, Heddle N, J K. An index of scientific quality for health reporters in the lay press. Journal of Clinical Epidemiology. 1993;46(9):987-1001.
94. Moynihan R. Tipsheet for reporting on drugs, devices and medical technologies,. http://wwwcommonwealthfundorg/Publications/Other/2004/Sep/Tipsheet--For-Reporting-on-Drugs--Devices-and-Medical-Technologiesaspx. 2004.
95. Guyatt GH, Ray J, Gibson N, Cook DJ, Ashpole B, Baines C, et al. A Journalist's Guide for Health Stories. American Medical Writers Association Journal. 2011;14(1):32-41.
96. Charnock D. The DISCERN Handbook: Quality Criteria for Consumer Health Information on Treatment Choices. Ltd, ORMP, editor. 1998.
97. Oxman AD, Chalmers I, Austvoll-Dahlgren A. Key Concepts for assessing claims about treatment effects and making well-informed treatment choices. F1000 Research. 2018.
98. Juri P. The Delphi method: substance, context, a critique and an annotated bibliography.1971.
99. Hasson F, Keeney S, McKenna H. Research Guidelines for the Delphi Survey Technique. Journal of Advanced Nursing. 2000;32(4).
100. Leonard M. Rasch Promises: a Layman's Guide to the Rasch Method of Item Analysis. Educational Research. 1980;22(3):188-92.

101. Austvoll-Dahlgren A, Semakula D, Nsangi A, Oxman AD, Chalmers I, Rosenbaum S, et al. Measuring ability to assess claims about treatment effects: the development of the “Claim Evaluation Tools”. BMJ Open. 2016;6(e013184).
102. Angoff WH. Scales, norms, and equivalent scores. In: Thorndike RL (ed.). Educational Measurement Washington DC. 1971:514-5.
103. Nedelsky L. Absolute grading standards for objective tests. Education and Psycholgical Measurement Journal. 1954;14(1):3-19.
104. Rosenbaum S. Improving the User Experience of Evidence: A Design Approach to Evidence-Informed Health Care. . Oslo: Arkitektur- og designhøgskolen i Oslo. 2010;https://brage.bibsys.no/xmlui//handle/11250/93062. Accessed 30 October 2018.
105. IDEO. Human Centered Design Toolkit. 1st Ed San Francisco:IDEO. 2011.
106. Abras C, Maloney-Krichmar D, Preece J. User-Centered Design. In Bainbridge, W Encyclopedia of Human-Computer Interaction Thousand Oaks: Sage Publi-cations, 2004. 2004.
107. Hillstrom K, Hillstrom LC, (Editors). Brainstorming. In: Encyclopedia of small business, Farmington Hills: Gale Cengage. 2002;1.
108. de Bono E. Lateral Thinking. A textbook of Creativity. The Penguin Group. 1970.
109. Morville P. User Experience Design: honey comb model. Sematic Studios LLC. 2004(Updated June 21, 2004).
110. Nsangi A, Semakula D, Glenton C, Lewin S, Oxman AD, Oxman M, et al. Resources to teach primary school children in low-income countries to assess claims about treatment effects; Protocol for a process evaluation. IHC working paper, http://wwwinformedhealthchoicesorg/wp-content/uploads/2016/08/Claim-cut-off-IHC-Working-Paper-2017-01-09pdf. 2016;Accessed 30 October 2018.
111. Lewin S, Glenton C, Munthe-Kaas H, Carlsen B, Colvin CJ, Gulmezoglu M, et al. Using qualitative evidence in decision making for health and social interventions: an approach to assess confidence in findings from qualitative evidence syntheses (GRADE-CERQual). . Plos Med 2015;12(10:):e1001895.
112. The Informed Healthcare Choices Group. Teachers' Guide for the Health Choices Book: Learning to think carefully when making choices about treatments. A health science book. Norwegian Institute of Public Health. 2016.
113. Nosich GM. Learning to Think Things Through: A Guide to Critical Thinking Across the Curriculum. Pearson Prentice Hall. 2009;3rd Edition.
114. Howieson C, Iannelli C. The effects of low attainment on young people's outcomes at age 22–23 in Scotland. British Educational Research Journal. 2008;34(2):269-90.
115. Bruner JS. The Process of Education. Cambride, MA: Harvard University Press. 1960.
116. Custers EJFM. Long-term retention of basic science knowledge: a review study. . Adv Hea Sci Educ 2010;15:109-28.
117. Ritter F, Baxter G, JW K, Srinivasmurthy S. Learning and retention. In Lee JD, Kirlik A, Dainoff MJ, editors. The Oxford handbook of cognitive engineering Oxford: Oxford University Press; 2013. 2013.
118. Tamim RM, Borokhovski E, Pickup D, Bernard RM. Large-scale, government-supported educational tablet initiatives. Burnaby, Canada: Commonwealth of Learning, 2015. 2015.
119. Kaplan B, Duchon D. "Combining Qualitative and Quantitative Methods in Information Systems Research: A Case Study”. MIS Quarterly. 1988;December 1988:571-86.
120. Slavin R, Madden NA. Measures inherent to treatments in program effectiveness reviews. Journal Res Educ Eff. 2011;4(370-80).
121. Martinez ME. Cognition and the question of test item format. Educational Psychology. 1999;34:207-18.
122. Stanger-Hall KF. Multiple-choice exams: an obstacle for higher-level thinking in introductory science classes. CBE life sciences education. 2012;11(3):294-306.

123. Simkin MG, Kuechler WL. Multiple choice tests and student understanding: what is the connection? Decision Science Journal Innovative Education. 2005;3:73-97.
124. Zeidner M. Essay versus multiple-choice type classroom exams: the student perspective. Journal of Educ Res. 1987;80:352-8.
125. Scouller K. The influence of assessment method on students' learning approaches: multiple choice question examination versus assignment essay. Higher Education. 1998;35:453-72.
126. Crane S, Broome ME. Understanding Ethical Issues of Research Participation From the Perspective of Participating Children and Adolescents: A Systematic Review. Worldviews on Evidence-Based Nursing. 2017;14(3):200-9.
127. Burke TM, Abramovitch R, Zlotkin S. Children's understanding of the risks and benefits associated with research. . Journal of Medical Ethics. 2005;31(12):715-20.
128. Cheah PY, Parker M. Consent and assent in paediatric research in low-income settings. BMC Medical Ethics. 2014;15(1):22.
129. Cohn JM, Ginsburg KR, Kassam-Adams N, Fein JA. dolescent decisional autonomy regarding participation in an emergency department youth violence interview. American Journal of Bioethics. 2005;5(5):70-4.
130. Chambers CT, Johnston C. Developmental differences in children's use of rating scales. Journal of Pediatric Psychology. 2002;27:27-36.
131. The Informed Health Choices group. Guide for piloting the Informed Health Choices (IHC) school resources. Informed Health Choices Working Paper, Norwegian Institute of Public Health, 2017 https://wwwinformedhealthchoicesorg/wp-content/uploads/2016/08/GUIDE-for-piloting-IHC-primary-school-resourcespdf 2017.
132. Oxman AD, Chalmers I, Austvoll-Dahlgren A. Key Concepts for assessing claims about treatment effects and making well-informed treatment choices [version 2; peer review: 3 approved]. F1000Research, https://f1000researchcom/articles/7-1784/v2. 2019;7(1784).
133. Lund HM, Rekkavik ME, Voll E, Mathisen PE. Teaching critical thinking about health claims: market analysis for Norwegian primary and lower secondary school. Informed Health Choices Working Paper, 2018. 2018.
134. McEwan PJ. Improving learning in primary schools of developing countries: a meta analysis of randomised experiments. Rev Edu Res. 2015;85:385-94.
135. Snilstveit B, Stevenson J, Phillips D, et al. Interventions for improving learning outcomes and access to education in low- and middle-income countries: a systematic review. London: International Initiative for Impact Evaluation, 2015. 2015;http://www.3ieimpact.org/en/evidence/systematic-reviews/details/259/ (accessed 11 Feb 2019).

21 APPENDICES

21.1 User Testing Interview Guide- Teachers' Resources (Appendix 1)

Informed Healthcare Choices – user test interview guide (Teacher's resources)	
Test person no.:	
What was tested:	TEACHERS' GUIDE
School:	
Date:	
Interviewer/Observer:	
What was tested (Chapter/Example):	CHAPTER ONE
Audio or video recording?	

Introduction, information, consent form, etc. (Introduce yourself, project and go through the informed consent form)

Turn on audio/video recording

Feedback on materials

1. Participant think-aloud (walkthrough each page of teacher's guide)

Where	Observer notes
Give teacher's both teacher's guide and copy of student book chapter. Explain that these are **sketches** of the teacher's guide and children's workbook. (Don't explain the embedding, but they should have a copy of the children's workbook so they understand what the materials consist of.)	
Front page	
Page 2 – Table of contents	
Page 3 – Picture with text "Dear teacher…"	
Pages 4 -5 - Introduction (the project)	
Page 6 - 7 –Teacher preparation for this chapter, lesson plan, lesson goals	
Page 8 - First intro pages from children's workbook • **Prompt:** this is the first page they see that is an embedded children's workbook page. Observe	

how easily they pick up on that, or if they don't notice, when do they understand this later. • Is the first time they see red items? Observe how they respond to them, if they seem to find them intuitive or confusing.	
Page 9 – Character intro	
Page 10-11 – Chapter 1 title and picture of schoolyard	
Page 12 - 13	
Page 14 – 15 • **Prompt:** this is the first time they see a glossary term	
Page 16 - 17	
Page 18 (last story page)	
Page 20 - Exercise A and Exercise B	
Page 21 – Activity • **Prompt:** elicit feedback **on the activity itself as well as the way it is presented**/formatted here.	
Page 22 .- Back cover	

B. Feedback (experience) on materials.

Question(s)	**Draft Responses**
1. What was your first impression about the teacher's guide? • **Prompt!** Can you remember what your first impression was of the guide when you first got the material? • **Prompt!** Find out if this particular guide was any different from the usual guides they use. If so, how different?	
2. Would you consider the teacher's guide difficult or easy to understand? • **Prompt**! Can you think of any steps or information that would simplify the guide even further? • **Prompt**! Can you point out sections or steps you consider difficult that could be taken out?	
3. Did you find the teachers' guide easy or difficult to use? In particular: • **Language used?** • **Readable type/layout?** • **Page size?** • **Design of instructions on the pages?**	

4. Do you think this guide would be useful to you when preparing for a lesson? • **Prompt!** Ask why it would or would not be?	
5. Can you say something about your experience of the trustworthiness of the material you have just read in the teacher's guide? • **Prompt**! Ask the participant if they can trust it? If they do or don't, what is the reason for their decision.	
6. Do you think this material was developed for a school like your own? • **Prompt!** Find out if it feels foreign or local, and why/where • how that affects their experience of the materials	

7. Do you have any suggestions for a title for the material in the children's book? • **Prompt**! You can let the participant know that we can get back to them later for a title suggestion?	
8. Can you sum up, three things that you liked and three you didn't like that we should carefully think about when re-designing the guide? **Prompt**! Anything they think we should definitely keep. **Prompt!** Anything they strongly feel should come out?	
9. Any other comments?	

C. Brief questions about the participant.

Question(s)	Draft Responses
10. What is your highest level of education?	
11. How long have you practiced as a teacher? • **Prompt!** Find out how many years they have been in ACTIVE practice.	
12. What subjects do you teach? • **Prompt**! If they teach all the four subjects find out their primary subject of specialization.	
13. What is your current work position? • **Prompt!** Find out if they also have any additional responsibilities at school. For example; If they are the School Director of Studies.	

D. Feedback on our session.

Question(s)	Draft Responses
1. How has this interview been conducted?	

End of session – stop and save recording

Thank the participant.

E. Immediate discussion after the session

Question(s)	**Draft Responses**
Points of discussion. • **Prompt!** (What stood out?) • **Learning points**	

F. Transcribing and Analysis

Person responsible: Lead Researcher.

21.2 User Testing Interview Guide- Children's Resources (Appendix 2)

Informed Healthcare Choices – user test interview guide (Children's resources)	
Test child's no:	
What was tested:	CHILDREN'S BOOK
Chapter tested:	CHAPTER ONE
School:	
Gender	
Age	
Education Level (Class)	
Favourite subject	
Date:	
Interviewer/Observer:	
What was tested (Chapter/Example):	

Feedback on materials

1. Participant think-aloud (walkthrough each page of teacher's guide)

Where	Observer notes
Give the child a copy of the student book chapter. Explain that these are **sketches** of the children's workbook.	
Before reading the book! **Prompt!** Ask what they think of this title, "The Fair and Compare Checklist" **Prompt**! Ask if they want to know what the title means now or they want to figure it out as they read the book?	
Before reading the book! **Prompt!** Ask what they would call this book, or what they would choose to call it from this list: workbook, book, handbook, textbook, school book, reader learning book.	
Before reading the book! **Prompt!** Remind them to circle any difficult word they come across.	

From start to finish, observe how long it takes them to read the book.	**TIME THEM (But do not let the child see you timing them- they might panic)**
Front page "This book belongs to"	
Page 2 – This book will teach you about health, treatments and claims.	
Page 3 – Pictures with character introduction… **Prompt!** What do they think of the characters and their names?	
Pages 4 -5 – Chapter 1 title and picture of schoolyard	
Page 6 - 7 _ John and Julie's first lesson of the day is science, with teacher Andy… **Prompt!** Familiarity with reading comics- understand which direction to read?	

Page 8, 9, 11,12 - • **Prompt!** Ask what they think about the definitions at the bottom of the page. Are they helpful?	
Page 13, 14 – • **Prompt!** Observe how long they take to do the exercise. • Was it fun? • How easy was the exercise	
Page 15 – **Observation:** How long it takes to complete the activity? **Prompt!** • Was it fun? • How easy was the exercise	
At the end- **Prompt!** Ask what they think they will learn from the book?	
At the end- **Prompt!** Ask for suggestions for a title for this book.	

Question(s)	Draft Responses

14. Do you think this material was developed for a school like your own? • **Prompt!** Find out if it feels foreign or local?	
15. What can we improve upon in the book? • **Prompt!** Ask for general suggestions for making it better.	
16. Any other comments?	

B. Feedback on our session.

Question(s)	Draft Responses
17. How has this interview been conducted? **Prompt!** What can we make better?	

End of session – stop and save recording

Thank the participant.

C. Immediate discussion after the session

Question(s)	Draft Responses
Points of discussion. • **Prompt!** (What stood out- main findings?) • **Learning points**	

D. Transcribing and Analysis

Person responsible: Research Team.

21.3 Results for children, parents and Teachers (Appendix 3)

Concept	Control	Intervention	Adjusted difference* (95% CI)	Difference per 100		Odds ratio (95% CI)
				← Favours control	Favours → intervention	
PASSING†						
Children *Primary school resources*	26.8%	69.0%	**49.8%** (43.8% to 54.6%)		49 more children per 100	9.3 (6.6 to 13.2)
Parents *Podcast*	37.7%	70.5%	**34.0%** (26.2% to 40.7%)		34 more parents per 100	3.9 (2.8 to 5.6)
Teachers *Primary school resources*	86.6%	97.6%	**11.3%** (4.0% to 13.0%)		11 more teachers per 100	7.2 (1.5 to 35.3)
MASTERY‡						
Children *Primary school resources*	0.9%	18.6%	**18.0%** (17.5% to 18.2%)		18 more children per 100	35.3 (20.6 to 60.7)
Parents *Podcast*	6.2%	31.6%	**26.0%** (15.2% to 39.1%)		26 more parents per 100	7.0 (4.0 to 12.1)
Teachers *Primary school resources*	14.9%	71.8%	**56.7%** (37.3% to 70.4%)		57 more teachers per 100	14.4 (6.2 to 33.1)

22 PAPERS I-IV

I. **<u>Nsangi A</u>**, Semakula, D, Rosenbaum SE, Oxman AD, Oxman M, Morelli A, Austvoll-Dahlgren A, Kaseje M, Mugisha M, Uwitonze A, Glenton C, Lewin A, Fretheim A, Sewankambo NK.*Development of the informed health choices resources in four countries to teach primary school children to assess claims about treatment effects: a qualitative study employing a user-centred approach***.** Pilot Feasibility Stud **6,** 18 (2020). https://doi.org/10.1186/s40814-020-00565-6

II. **<u>Nsangi A</u>**, Semakula D, Oxman AD, Austvoll-Dahlgren A, Oxman M, Rosenbaum ES, Morelli A, Glenton C, Lewin S, Kaseje M, Chalmers I, Fretheim A, Ding Y, Sewankambo NK. *Effects of the Informed Health Choices primary school intervention on the ability of children in Uganda to assess the reliability of claims about treatment effects: a cluster-randomised controlled trial.* Lancet 2017; doi.org/10.1016/ S0140-6736(17)31226-6

III. **<u>Nsangi A</u>**, Semakula D, Glenton C, Lewin S, Oxman AD, Oxman M, Rosenbaum S, Austvoll-Dahlgren A, Nyirazinyoye L, Kaseje M, Rose CJ, Fretheim A, Sewankambo NK. *Informed Health Choices intervention to teach primary school children in low-income countries to assess claims about treatment effects: process evaluation.* BMJ Open 2019;9:e030787. doi: 10.1136/bmjopen-2019-030787

IV. **<u>Nsangi A</u>**, Semakula D, Oxman AD, Austvoll-Dahlgren A, Oxman M, Rosenbaum SE, Morelli A, Glenton C, Lewin A, Kaseje M, Chalmers I, Ding Y, Fretheim A, Sewankambo NK. *Effects of the Informed Health Choices primary school intervention on the ability of children in Uganda to assess the reliability of claims about treatment effects, 1-year follow-up: a cluster-randomised trial.* Trials 21**,** 27 (2020) doi:10.1186/s13063-019-3960-9

I

Nsangi *et al. Pilot and Feasibility Studies* (2020) 6:18
https://doi.org/10.1186/s40814-020-00565-6

Pilot and Feasibility Studies

RESEARCH Open Access

Development of the informed health choices resources in four countries to teach primary school children to assess claims about treatment effects: a qualitative study employing a user-centred approach

Allen Nsangi[1,2], Daniel Semakula[1,2], Sarah E. Rosenbaum[3], Andrew David Oxman[2,3*], Matt Oxman[3], Angela Morelli[4], Astrid Austvoll-Dahlgren[3], Margaret Kaseje[5], Michael Mugisha[6], Anne-Marie Uwitonze[6], Claire Glenton[3], Simon Lewin[3,7], Atle Fretheim[2,3] and Nelson Kaulukusi Sewankambo[1]

Abstract

Background: People of all ages are flooded with health claims about treatment effects (benefits and harms of treatments). Many of these are not reliable, and many people lack skills to assess their reliability. Primary school is the ideal time to begin to teach these skills, to lay a foundation for continued learning and enable children to make well-informed health choices, as they grow older. However, these skills are rarely being taught and yet there are no rigorously developed and evaluated resources for teaching these skills.

Objectives: To develop the Informed Health Choices (IHC) resources (for learning and teaching people to assess claims about the effects of treatments) for primary school children and teachers.

Methods: We prototyped, piloted, and user-tested resources in four settings that included Uganda, Kenya, Rwanda, and Norway. We employed a user-centred approach to designing IHC resources which entailed multiple iterative cycles of development (determining content scope, generating ideas, prototyping, testing, analysing and refining) based on continuous close collaboration with teachers and children.

Results: We identified 24 Key Concepts that are important for children to learn. We developed a comic book and a separate exercise book to introduce and explain the Key Concepts to the children, combining lessons with exercises and classroom activities. We developed a teachers' guide to supplement the resources for children.

Conclusion: By employing a user-centred approach to designing resources to teach primary children to think critically about treatment claims and choices, we developed learning resources that end users experienced as useful, easy to use and well-suited to use in diverse classroom settings.

Keywords: User-centred design, User-testing, User experience, Pilot study, Critical thinking, Critical appraisal, Teaching, Education

* Correspondence: oxman@online.no
[2]University of Oslo, Oslo, Norway
[3]Centre for Informed Health Choices, Norwegian Institute of Public Health, Postboks 222 Skøyen, 0213 Oslo, Norway
Full list of author information is available at the end of the article

Article summary

Strengths and limitations of this study

Strengths

- We used a user-centered design approach with a multi-disciplinary team.
- We engaged end-users in the entire development process from brainstorming to piloting.
- Non stringent grant conditions permitted ample time to generate and prototype ideas and then iteratively design the resources.

Limitations

- Time constraints in trying to synchronise the design schedule with the already busy school schedule

Summary box

What is already known:

- There is an information overload regarding unsubstantiated claims of benefits and harms of treatments
- People generally lack the skills to assess the reliability of treatment claims
- Lack of resources to teach critical thinking particulary appraising treatment claims in primary schools in both low and high-income countries.

What are the new findings:

- Use of a user-centered design approach to design resources
- Benefits of multi-stake holder collaboration in the design process

How might it impact on clinical practice in the foreseeable future?

- We designed useful, understandable and transferable resources to teach critical thinking that children and teachers found relevant and easy to use in their particular contexts.

Background

People of all ages, in low- and high-income countries, are flooded with both reliable and unreliable information about how to care for their health, including claims about the benefits and harms of treatments (any action intended to improve health) [1]. Unreliable claims come from many sources, including experts, advertisements and family [2]. People's beliefs in unproven claims about treatments can lead to harm and waste [2]. Although this problem is global, people with fewer resources to spend on unnecessary treatments are disproportionately affected.

Many studies have found that people's ability to understand and assess health information is often lacking [1, 3–5], although there are limitations in how this has been measured [6]. The Informed Health Choices project aims to enable people to assess claims about the effects of treatments, beginning with primary school children.

Why target primary school children?

Research has suggested that children between the ages of 10 and 12 are capable of learning critical appraisal skills [7], and teaching these basic skills is already part of the curricula in some countries [8]. It is possible to reach a large segment of the population before they drop out of school, as many do after primary level in low-income countries [9–11]. Finally, teaching children to assess information about treatment effects can lay a foundation for them to make informed health decisions when they grow older, as patients, future health professionals, policymakers and citizens.

A recent overview of six systematic reviews on education interventions in under resourced countries included 227 studies in total, but none of these studies addressed health or science literacy, or critical thinking more broadly [12]. Systematic reviews of teaching children critical appraisal skills in health also have not found studies of strategies for teaching these skills to primary school children in both low and high income settings [6, 13].

We developed the Informed Health Choices (IHC) primary school resources to help children begin to learn critical appraisal skills required to assess benefits and harms of treatments. Our objective was to design resources that children and teachers experienced as useful, easy to use, understandable, credible, desirable, and well-suited in classroom settings. In this article, we describe the development of these resources.

Methods

Researchers in Norway, the United Kingdom, Uganda, Kenya, and Rwanda collaborated to develop and evaluate learning resources for school children and their parents in 2013 to 2017. This included development of a podcast for parents [14]; development of the CLAIM Evaluation Tool for measuring people's ability to assess treatment claims [15]; a randomised trial of the effects of using the (IHC) primary school resources [16]; a randomised trial of the effects of listening to the podcast [17]; and a process evaluation [18].

Participants and setting

While most of the piloting, user testing, and feedback took place in Uganda, we wanted to create resources that could also be used in other countries. Therefore, we also carried out piloting and user-testing of Version 2 of the resources in two East African countries (Rwanda and Kenya) and in one high-income country (Norway).

For pilot testing, we recruited schools that were geographically accessible to our team, taught in English, and were willing to make time. We contacted head teachers, who identified science teachers and classes of children who were prepared to pilot the resources. To recruit user-test participants, we used purposeful sampling to include year 5 students (10 to 12-year-olds) and their teachers. Table 1 describes the participants, and (Additional file 1) describes which participants we included in each step of the development work.

Developing the resources

We employed a user-centred approach to designing the IHC primary school resources [19–22]. User-centred design is characterised by multiple iterative cycles of development (Fig. 1).

Our starting point for developing these learning resources was to create a list of 32 Key Concepts that people need to understand and be able to apply to assess claims about treatment effects and make informed health choices [15]. A network of teachers in Uganda assessed the relevance of these concepts for primary school children during the prioritisation process and determined that 24 of these concepts were relevant to primary school children in Uganda [23].

Idea generation and prototyping

We used "creative thinking" in the idea generation and prototyping process. Creative thinking focuses on exploring ideas, generating possibilities and looking for many options [24]. This contrasts with critical assessment, which focuses on analysis, figuring out the answer and eliminating incorrect options. Both types of thinking were necessary for generating appropriate options for the resources we developed.

We needed to bring stakeholders and end users as close as possible into all phases of the work. This was particularly important since none of us belonged to the end user groups we were developing resources for (primary school children and their teachers). We included teachers as close collaborators through brainstorming [25] and prototyping workshops, and sought early feedback from children through workshops and school visits. We conducted multiple workshops in Uganda and Norway with the IHC research team and a network of teachers in Uganda [23]. These workshops resulted in ideas and insights about the context and stakeholders, sketches, and prototypes. We selected and developed ideas that we thought had the most potential to create new prototypes. These prototypes formed the basis for the next phases of pilot testing and user-testing.

Pilot testing and user-testing

We pilot tested early prototypes in workshops with teachers and children and through school visits in Uganda and Norway, using participatory observation to facilitate participants' engagement. We piloted later, more complete prototypes (Version 1 in Uganda, and

Table 1 Participants

Participants	Description
Researchers, teachers and journalists from several countries	The initial brainstorming session at the kick-off meeting for the project included 18 people from Indonesia, Nepal, Norway, Uganda, and the United Kingdom with various backgrounds, including teachers, journalists, medical doctors, information designers, anthropologists, public health specialists, and health service researchers.
A national advisory board in Uganda	The advisory board for the project included fifteen members (2 women and 13 men) representing various stakeholders, including the Ministry of Education, Ministry of Health, and Ministry of Gender, Labour and Social Development (which is responsible for children's affairs in Uganda), and representatives from civil society and local government.
A teachers' network in Uganda	The teachers' network included 24 Ugandan primary school teachers (10 women and 14 men) in active practice from both rural and urban schools that were either government or privately owned [19].
Schools in Uganda	Of the five schools that participated in both phases of the development process (pilot and user-testing), four were government (public) schools and one was a private school. One of the government schools was one of the biggest schools in the country, with a teacher-student ratio of 1:250. The other three government schools were of typical size, with a teacher-student ratio of 1:120. The private school was small, with a teacher-student ratio of 1:35, in comparison to the average Uganda school with a teacher student ration of 1:70. For logistic purposes (travel by the investigators), three of the schools that participated were located in the Kampala urban area and two were in the semi-urban area surrounding Kampala. All of the schools were poorly equipped. Lessons were in English, although English was not the primary language spoken at home for most of the children. All of the classes were year-5, for which the official starting age is 10.
A school in Kenya	The school in Kenya was a government school with about 400 children attending year-1 to year-8 classes. The year-5 children were mostly between 10 and 14 years old.
A school in Rwanda	The school in Rwanda was a government (public) primary and secondary school with over 3000 children. The language of instruction was English and the age range for year-5 children was 10 to 15 years old.
Children in Norway	A convenience sample of four 12-year-old girls who knew each other, from a nearby school participated in piloting a series of eight games together with the research team, partly in Norwegian and partly in English.
A school in Norway	The school in Norway was a private international school, with 18 children in each class. It was well equipped. Lessons were in English, although English was not the primary language spoken at home for most of the children. The three classes were year-7, for which the typical starting age is 11.

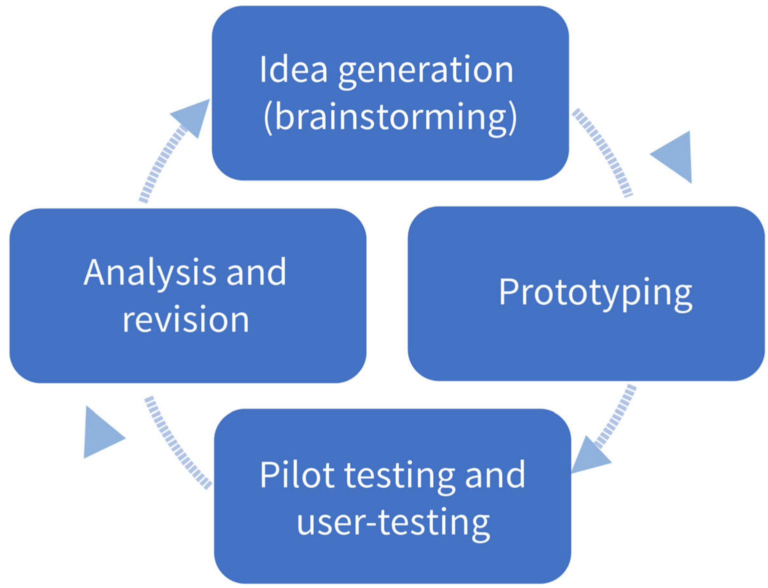


Fig. 1 User-centred design development in multiple iterative cycles

Version 2 in Uganda, Rwanda, Kenya and Norway) using non-participatory observations of the classroom lessons to explore how teachers and children used these resources. We used a structured form to record observations (Additional file 2), as well as video and still photography.

We also carried out user-test interviews with individual children and teachers to explore their experience when interacting with our resources [19]. User-testing originated from human computer interaction, where effectiveness and efficiency of a product is measured in relation to personal satisfaction of the individual using the product. We used a qualitative approach, building on Rosenbaum's adaptation of Peter Moville's honeycomb framework of user experience [19–22, 25–27] to develop the interview guides. We focused on six facets of the users' experiences: usefulness, ease of use, understandability, credibility, desirability, and identification (Table 2) [19].

Table 2 Six facets from the honeycomb framework

Facet	Description
Usefulness	Does this product have practical value for this user?
Usability	How easy and satisfying is this product to use?
Understandability	Does the user recognise what the product is and do they understand the content? (own subjective experience of understanding)
Credibility	Is it trustworthy?
Desirability	Is it something the user wants - has a positive emotional response to?
Identification	Does the user feel the product is for" someone like me" or is it alienating/foreign-feeling? (e.g. age, gender, culture–appropriate)

Analysis and revisions

We used a framework analysis approach to guide data collection and analysis. We entered observations from the pilot testing and feedback from the user-testing into a spreadsheet after each round of testing. Between two and five researchers from the IHC working group independently coded each observation based on the importance of the finding (Table 3) and its implications for changes to the learning resources.

The coding was combined in a single spreadsheet, discussed, and a consensus was reached. Based on these findings, we generated a list of potential problems and suggestions for changes. We discussed major problems and brainstormed solutions to those problems with the rest of the IHC working group. After agreeing on the changes needed, we created new prototypes to be piloted and user-tested.

We did not collect or analyse any quantitative data.

A timeline showing the development process, beginning with prioritisation of the Key Concepts is shown in (Fig. 2), and each step is summarised in (Additional file 1).

Patient and public involvement statement

End users such as teachers on the network, policy makers on the advisory panels and primary school children participated in the development process by providing structured feedback of the resources at various iteration stages.

Results

Prioritising key concepts

We started with the list of 32 Key Concepts identified at the beginning of this project [15]. However, although 24

Table 3 Coding of the importance of observations and feedback

Code	Description
Very important negative finding ("show stopper")	A problem that we should address for the resources to be effective
Important negative finding	A problem that we should probably address for part of the resources to be effective
Negative finding	A problem that we can easily address and probably will not prevent the resources from being effective
Very important positive finding	Praise that probably should inspire changes
Important positive finding	Praise that maybe should inspire changes
Positive finding	Praise that probably should not inspire changes
Very important constructive finding	A suggestion that probably should inspire changes
Important constructive finding	A suggestion that maybe should inspire changes
Constructive finding	A suggestion that probably should not inspire changes

of these concepts were deemed relevant for primary school children, teaching all 24 concepts proved to be too much to learn in a school term. The early prototypes we created had too many concepts per lesson and took too long to teach in a normal school hour (40 min). We also observed that the teachers needed time to repeat material from previous lessons. We therefore reduced the number of concepts addressed in the final version of the resources to 12 (Table 4), as described in (Additional file 3). The other 12 concepts could be taught in a subsequent school term.

Review, idea generation and exploratory prototypes

This phase lasted two years and was highly exploratory. In addition to the workshops and prototype development described below, we also engaged regularly with the teachers' network and the Uganda National Advisory board.

Reviewing existing resources

We collected ideas from our own experiences teaching critical appraisal to children [28] and adults (including health professionals, policymakers, journalists, and patients), a systematic review of interactive resources for teaching critical appraisal skills to consumers [29], and searching the TES database and other sources such as google scholar for relevant resources.

We had a series of brainstorming sessions with members of the research team, informed by the resources that we found and workshops that we conducted with teachers and other researchers. In October 2015, we organised an international workshop with others interested in helping people to assess claims about treatments where a variety of resources was discussed. This workshop led to the development of the Critical thinking and Appraisal Resource Library (CARL) [30]. The Critical thinking and Appraisal Resource Library (CARL) is a platform to collect

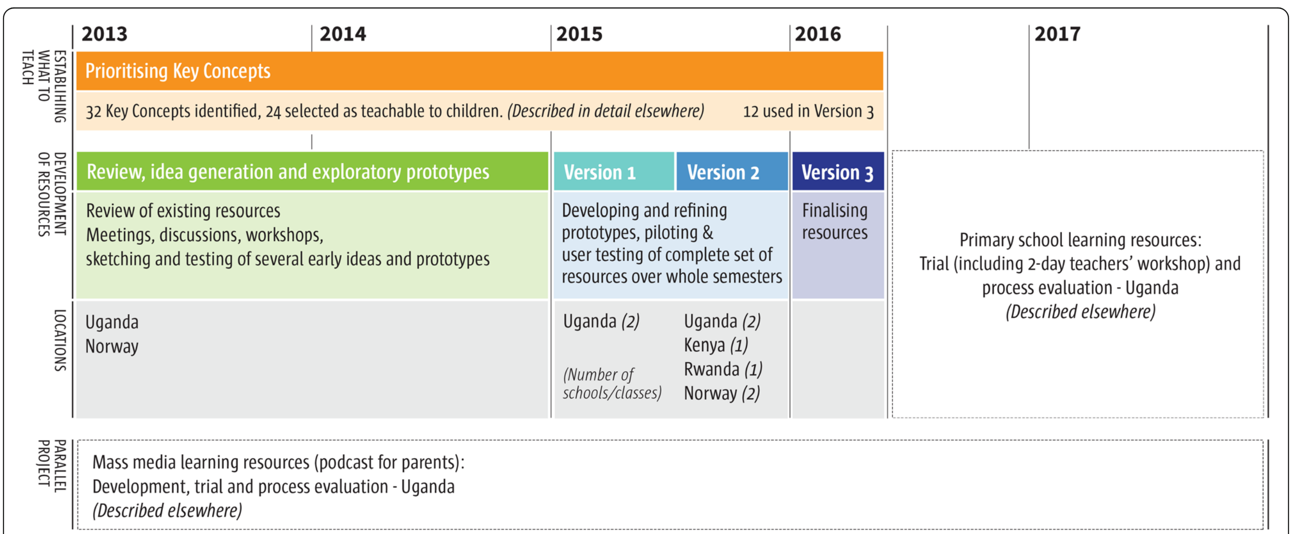


Fig. 2 Development timeline

Table 4 Key Concepts that are relevant for primary school children

Key Concepts taught in The Health Choices Book
CLAIMS: ARE THEY JUSTIFIED?
• Treatments may be harmful
• Personal experiences or anecdotes (stories) are an unreliable basis for assessing the effects of most treatments
• Widely used treatments or treatments that have been used for a long time are not necessarily beneficial or safe
• New, brand-named, or more expensive treatments may not be better than available alternatives
• Opinions of experts or authorities do not alone provide a reliable basis for deciding on the benefits and harms of treatments
• Conflicting interests may result in misleading claims about the effects of treatments
COMPARISONS: ARE THEY FAIR AND RELIABLE?
• Evaluating the effects of treatments requires appropriate comparisons
• Apart from the treatments being compared, the comparison groups need to be similar (i.e. 'like needs to be compared with like')
• If possible, people should not know which of the treatments being compared they are receiving
• Small studies in which few outcome events occur are usually not informative and the results may be misleading
• The results of single comparisons of treatments can be misleading
CHOICES: MAKING INFORMED HEALTH CHOICES
• Treatments usually have beneficial and harmful effects
Other Key Concepts prioritised for children
CLAIMS: ARE THEY JUSTIFIED?
• An outcome may be associated with a treatment, but not caused by the treatment
• Increasing the amount of a treatment does not necessarily increase the benefits of a treatment and may cause harm
• Hope or fear can lead to unrealistic expectations about the effects of treatments
• Beliefs about how treatments work are not reliable predictors of the actual effects of treatments
• Large, dramatic effects of treatments are rare
COMPARISONS: ARE THEY FAIR AND RELIABLE?
• People in the groups being compared need to be cared for similarly (apart from the treatments being compared)
•
• It is important to measure outcomes in *everyone* who was included in the treatment comparison groups
• Results for a selected group of people *within* fair comparisons can be misleading
• Reviews of treatment comparisons that do not use systematic methods can be misleading
• Well done systematic reviews often reveal a lack of relevant evidence, but they provide the best basis for making judgements about the certainty of the evidence
CHOICES: MAKING INFORMED HEALTH CHOICES
• Fair comparisons of treatments should measure outcomes that are important

and distribute freely-available learning resources intended to help people think critically about treatment claims.

Idea generation workshop with researchers, teachers and journalists

In this meeting, we generated a broad range of ideas, from holding science fairs to creating interactive videos. Some ideas we generated were: use of drama and storytelling, board and field games, getting children to run a trial over several months, building a collection of familiar examples, translating already existing resources into local languages, holding teacher training workshops.

The main challenges we identified included: the need to teach the teachers; developing resources that would work in schools without digital equipment and where languages other than English were spoken; finding time in the curriculum, and gaining buy-in from stakeholders (including teachers, parents, and policymakers).

We decided to focus the next step on developing interactive classroom games that could be carried out with simple readily available equipment, like blackboards.

Pilot testing games in classrooms

We developed presentation materials and prototypes for two games to be used in classrooms: tossing coins to explain the concept of 'chance' and a game involving comparing the effects of two different coloured candies to explain Key Concepts related to fair comparisons. Children worked together in small groups. We piloted the games in classes at three schools – one in Norway and two in Uganda, with numbers of children ranging from 30 to 129. We participated by taking the role of teachers.

The children clearly enjoyed these activities. They were engaged, asked relevant questions and came up with some of the concepts by themselves, like blinding. But the exercise tended to get out of hand when the children were required to work independently and discuss in small groups. This was a problem even in Norway, despite the smaller class size. The children also needed more structured materials and more facilitation than we had anticipated. Their understanding of the concept 'fair' was different than what we meant when talking about fair comparisons, which we referred to initially as "fair tests". One child said:

> "For the test to be a fair test, everyone should get a candy".

Despite being encouraged by the apparent ability of the children to understand many of the concepts, we also experienced first-hand that it could be challenging to explain the concepts correctly, even with semi-structured presentation materials. Teachers who were unfamiliar with the concepts would likely have even more difficulty.

Prototyping and pilot testing in Uganda and Norway

We conducted a prototyping workshop with 24 members of the teachers' network in Uganda, piloted a game at a school in Uganda and an international school in Norway, and piloted a series of eight games with four 12 year-old girls in Norway (Additional file 4).

We found that although some of the games appeared to be promising, several were still too complicated to carry out in large classrooms. We also still had not solved the problem of how to transfer our presentation role to a teacher who was unfamiliar with the concepts without relying on electronic equipment like PowerPoint or video.

We decided to produce a highly-structured narrative for presenting the Key Concepts, which the teacher and children would read together, as well as a guide for the teacher. We decided to make a narrative in the form of a comic book with game-like activities and individual exercises included. We developed five characters to build the story around: two school children, two professors and a parrot who made unreliable claims about treatments, in an unspecified setting that would look like a rural east African village. Our thinking was that the narrative and use of drawings would engage the children, make the Key Concepts easier to understand, and help them to retain what they learned [31, 32].

Despite many problems, there was enough enthusiasm for the comic format that we felt it had the potential to work in a Ugandan classroom. But it needed to be much simpler, and the explanation of each concept needed more space. Based on our findings and observations, we agreed to make the following changes in the next version:

- Rewrite and redraw the children's book with
  - A much simpler story, language, and drawings; and shorter chapters with larger text
  - No complicated comic language
  - Glossary explanations where terms first occur, with definitions translated to Luganda
  - Examples that were less likely to be misleading
  - Simplify all activities so they would not require extra resources, or require being outdoors
- Revise the teachers' guide by
  - Making it more like a recipe
  - Integrating the children's book in the teachers' guide to facilitate the lesson flow

We decided to produce the final version of the books in colour, but continued sketching prototypes in black-and-white.

The IHC primary school resources

We created three complete versions of the children's book and teachers' guide. The first version had 11 chapters (Fig. 3). We carried out pilot tests and user-testing at two schools in Uganda. Based on the users' experiences (Additional file 5), we made the following changes to the next version of the children's book:

- More emphasis on "critical thinking" rather than becoming a "junior researcher"
- Add a new first chapter that clarifies the purpose of the book, introduces some of the basic vocabulary in more depth ("health" "treatments", "effects" and "claims").

Fig. 3 Version 1 prototype of the IHC primary school resources

- Make usefulness more apparent by placing the story in the context of real life decision-making (e.g. the children in the book making a poor decision in the beginning and a more informed decision at the end)
- Adjust chapter content so that lessons could be fitted into 40-min periods
- Repeat learning goals from the previous chapter and introduce new characters at the beginning of each chapter
- Continue to simplify vocabulary; add a glossary in the back of the book
- Use a computer font instead of handwriting
- Add arrows to the comic cells to indicate reading direction
- More expressive and differentiated characters

We agreed on the following changes to the Teachers' guide:

- Introduce more structure
- Add more background information, both about the purpose of the resources and about the key concepts covered in each chapter
- Decrease the number of lesson goals in each chapter

We created Version 2 of the children's book and teachers' guide (Fig. 4), which had 10 chapters divided into two books. We carried out pilot tests and user-testing at schools in Uganda, Rwanda, Kenya, and Norway.

The most important problem that we identified was insufficient time to teach all the content included in Version 2. Based on the users' experiences (Additional file 5), we agreed to make the following changes in the next version of the children's book:

- Revise the CLAIM game and make it less demanding on the teacher to organise
- Introduce a glossary that explains all the new terms in the children's book
- Reduce the number of exercises at the end of each lesson
- Further simplify or remove chapters that were difficult for the children to understand, like chapter 8 on "careful summaries" (systematic reviews)

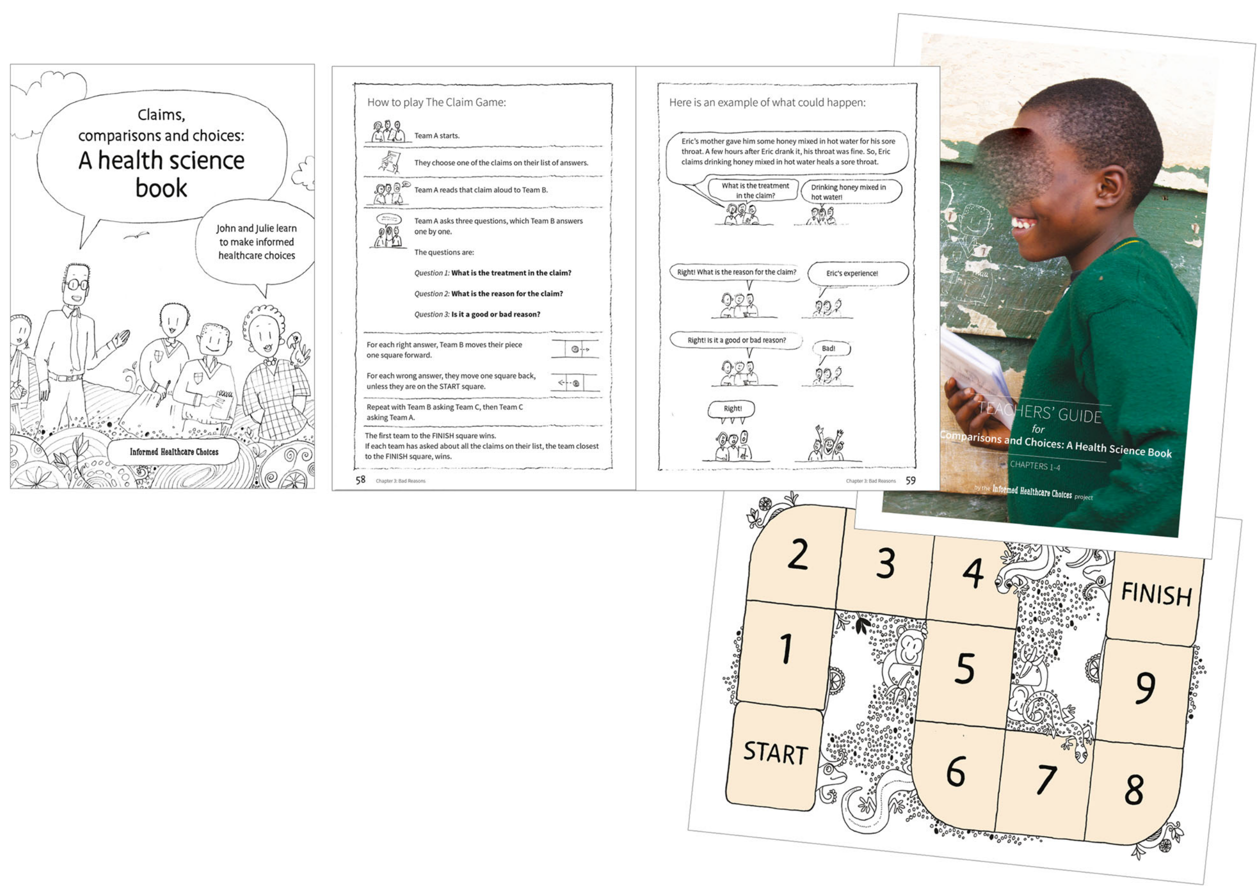


Fig. 4 Version 2 prototype of the IHC primary school resources

We agreed to make the following changes in the teachers' guide:

- Add more examples
- Revise and restructure the content and add a structured lesson plan

We created Version 3 of the children's book with 10 chapters, and a teachers' guide (Fig. 5). We also created a separate exercise book, a classroom poster of the key learning objectives (the 12 Key Concepts), and a set of activity cards for one of the chapters. These open access resources can be viewed or downloaded at http://www.informedhealthchoices.org/primary-school-resources/. The contents of the children's book and the teachers' guide are summarised in Table 5.

Discussion

While focussing on the six facets, (usefulness, ease of use, understandability, credibility, desirability and identification), of Rosenbaum's adaptation of Peter Moville's honeycomb frame work of user experience, this work highlights the following lessons for future studies designing educational materials;

Usefulness

Findings from the idea generation and exploratory prototypes phase of the project highlighted the need to clarify the usefulness of the resources for both teachers and children. Teachers' participating in the Uganda teachers' network workshop initially assumed that the purpose of the resources was to convey public health messages about the benefits of specific interventions, such as handwashing, healthy eating habits and exercise. Many of the ideas and prototypes generated at that workshop focused on communicating typical public health messages, rather than teaching children to think critically about health claims and choices.

There are several plausible explanations for this. These include that teaching is largely didactic in East Africa, in part due to large student-to-teacher ratios. This makes it difficult to use more interactive teaching strategies required to teach critical thinking skills [33]. Teaching critical thinking skills has not been a priority in primary school curricula or for evaluations of interventions to improve primary school education [12, 34, 35]. Previous public health interventions in schools have also tended to focus on promoting specific behaviours, rather than teaching children to think critically. This contrasts with

Fig. 5 Version 3 (final) of the IHC primary school resources

Table 5 Contents of the children's book and the teachers' guide

Children's book	Teachers' Guide
The Health Choices Book: Learning to Think Carefully about Treatments. A health science book for primary school children *Introduction* Lesson 1: Health, treatments and effects of treatments *John and Julie learn about CLAIMS about treatments* Lesson 2: Someone's experience using a treatment Lesson 3: Other bad bases for claims about treatments (Part 1) Lesson 4: Other bad bases for claims about treatments (Part 2) *John and Julie learn about COMPARISONS of treatments* Lesson 5: Comparisons of treatments Lesson 6: Fair comparisons of treatments Lesson 7: Big enough fair comparisons of treatments *John and Julie learn about CHOICES about treatments* Lesson 8: Advantages and disadvantages of a treatment *Review* Lesson 9: Review of what is most important to remember from this book Glossary	*Teacher' guide for the Health Choices Book* The teacher's guide includes an introduction to the project and the resources, and the following for each lesson, in addition to the embedded chapter from the children's book: • The objective of the lesson • A lesson preparation plan • A lesson plan • A list of materials that the teacher and children will need • A synopsis of the story • Keywords in the chapter • Review questions to ask the children after reading the story • Extra examples for illustrating the concepts • Background about examples used in the story • Teacher instructions for the classroom activity • Answers and explanations for the activity • Answers and explanations for the exercises • Background information, examples, and keyword definitions for teachers

our findings in Norway. Critical thinking was a priority for older children (in the International Baccalaureate IB programme) at the international school where we piloted the second version of the resources. However, the teachers there found that students entering the IB programme were not sufficiently prepared. They wanted to test our resources specifically to find out if they might help to address this problem that they had already identified.

Expectations of the children in response to early prototypes were different from those of the teachers. They assumed that the purpose of the resources was to help them do better in science and to learn to become scientists or health professionals.

We addressed these misunderstandings about why the resources are useful in several ways. We added introductions to both the children's book and the teachers' guide clarifying the purpose of the resources. These went through several iterations and we obtained feedback from teachers and children to ensure that the introductions clarified the purpose of the resources and why they are useful. We ensured that the examples we used would not be misunderstood and that they clearly illustrated how each Key Concept could be used to assess relevant claims and to make informed choices. We modified the structure of the book, and subsequently organised the Key Concepts (from six groups to three groups), to clarify and reinforce the purpose and usefulness of understanding and applying them.

When testing the first and second versions of the resources we found that teachers and most children found the resources useful and correctly understood their purpose by the end of the lessons. In addition to the above changes, we also developed a workshop for teachers to introduce them to the resources and to help ensure that they started out with a clear understanding of the purpose of the resources. The workshop is described in detail in another article [36].

Ease of use

We found that our initial ideas and prototypes were difficult to use, even in well-resourced schools with low student-to-teacher ratios. We also found that many of the Key Concepts were not well understood by the teachers. Frequently they went off script, making unsubstantiated claims themselves rather than helping the children learn how to assess claims. Using a comic book to introduce the Key Concepts solved the problem of ensuring that they were introduced and explained correctly. The illustrations facilitated engagement, understanding and made it easier for the children to read the text. This is consistent with previous research, which has shown that adding pictures to written language can increase attention, comprehension, and recall [32]. However, pictures can also be misunderstood and the feedback we received on the illustrations resulted in many changes - both specific and general. For example, feedback from several children resulted in changes to how Julie, one of the two children who are main characters in the comic book was portrayed. As one child remarked when asked about the drawings in an early version: "Julie looks like a rumour monger."

We also discovered important changes that were needed to make the comic book usable in Uganda. Many of the children were not familiar with reading comics and were confused about the order in which the frames should be read. They also were not familiar with speech and thought bubbles. To address this problem, we added arrows to the comic, showing the order in which frames should be read and explained speech and thought bubbles in the introduction.

Using a comic book to introduce the Key Concepts functioned well both in East Africa, where it is common for classes to read aloud and in Norway, where role playing was used when reading aloud in class. Based on our observations and interviews, we concluded that there were several ways of reading the book. Rather than recommending one of these, we provided the advice based on what we had observed.

Our observations and feedback from the teachers resulted in several changes to the teachers' guide to ensure that teachers found it useful. One change was to incorporate the children's book in the teachers' guide. This facilitated using the guide, which includes instructions and suggestions for the teachers, while reading the comic together with the children or doing the activities. Others included providing lesson plans, explanations written for the teachers, and extra examples that the teachers could use to illustrate the Key Concepts.

The most important problem that we found with the second version of the resources was insufficient time. Teachers struggled to get though the lessons in 40 min and, therefore, the children were often confused and had not learned some of the Key Concepts. To address this problem, we reduced the number of concepts that were included from 24 to 12 and we doubled the amount of time for each lesson. This required us to step back and acknowledge that we had made a classic mistake of trying to teach everything about a topic at once, thereby overloading both the children and the teachers with too much information. By recognising that the resources are just one cycle in a spiral curriculum [37], we could make this dramatic change. Resources for subsequent cycles can build on what was previously learned from these resources and reinforcing while introducing new concepts.

Other changes that we made to the resources to improve their usability included greatly simplifying the activities to ensure that they could easily be managed by a single teacher with many children and ensuring that the exercises could be done by the children without placing a substantial burden on the teacher.

Understandability

We discussed understanding the purpose of the resources in relation to its perceived usefulness and how that affected the extent to which teachers and children valued the resources. We also found substantial problems with understanding of the content. Many of the children read poorly and English was a second language for most. We found that words that we assumed 10 to 12-year-old children would understand, such as 'health', were new words for many children in East Africa. Although using a comic book with illustrations helped to improve understanding, we still needed to further simplify the language that we used explain terms. We addressed this by iteratively testing and rewriting the text, adding a glossary, adding translations of key terms to Luganda and Kiswahili, adding a list of new keywords used in each chapter, and adding explanations and translations of key terms to the text on the page where they were first used (Fig. 6). Together with teachers and children, we also generated a list of terms that were difficult for the children. We avoided using those terms if there was a good alternative or explained them.

Several changes to the teachers' guide were made to ensure their understanding, these included adding a background section to each chapter and extra information about the examples that we used (Fig. 7), in addition to the workshop for teachers noted above.

Credibility

Two problems that we identified were the use of magical elements in the first comic prototype and the inclusion of a talking parrot. We eliminated the former, but elected to keep the parrot for two reasons. First, although teachers were concerned that a talking animal would result in a loss of credibility amongst the children, none of the children perceived this as a problem. Second, the children responded very positively to the parrot, which both brought humour into the story and served as a source of claims. We did, however, review our use of the parrot to ensure that it was used consistently and that it was not included unnecessarily; e.g. repeating something that one of the other characters said.

Desirability

Many of our early ideas, which focused on games, were clearly not something that the teachers wanted. They were difficult to organise and to manage, especially in classes with large student-to-teacher ratios.

We found that the book was highly desirable both in East Africa and in Norway. This was, perhaps, not surprising in East Africa where the schools had few books. However, the children at the international school in Norway also were very positive about the book. They uniformly responded that they would prefer the book to a computer game. It is uncertain to what extent this was because they had been exposed to poorly designed learning games or because the book was well designed. Children in both settings had not previously been exposed to use of a comic book to teach science.

The rationale for using a narrative in the book to explain the Key Concepts is that people often make sense of their lives through stories they hear and share with others [31]. Providing information in a story may resonate with people who might struggle to understand abstract concepts. Furthermore, characters in the narrative can role model new behaviours, enhancing self-efficacy [38]. Evaluations of the effects of narrative interventions support their use. For example, evaluations of the use of narratives in the context of health promotion have found that narrative interventions improve knowledge about health-promoting behaviours and the behaviours themselves [31].

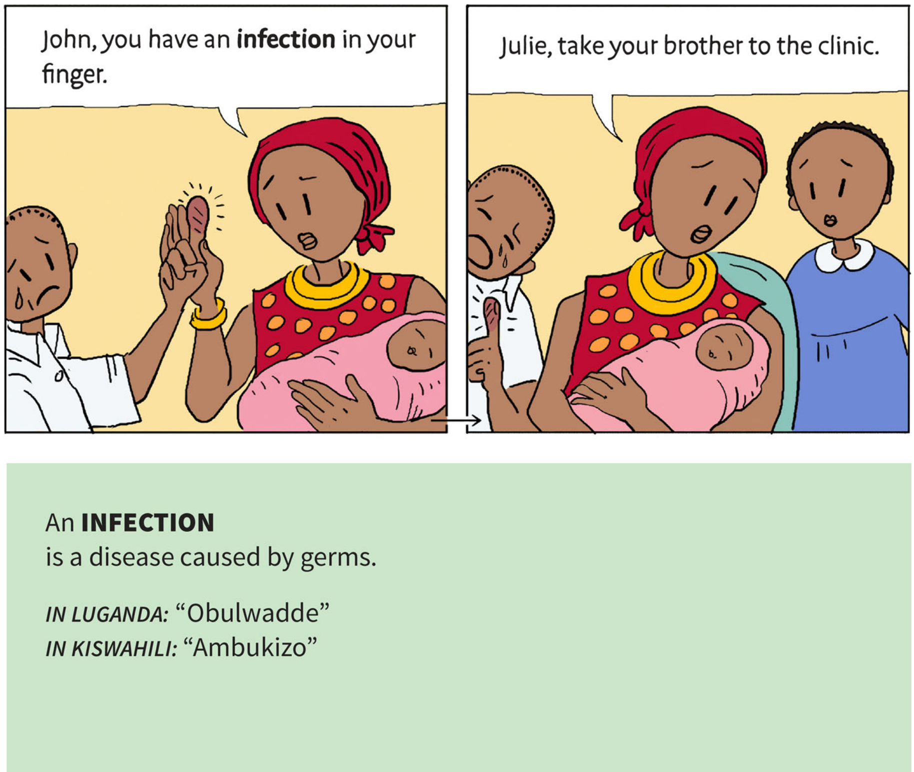


Fig. 6 Repeating keywords where they first appear in the text

Although we received consistent feedback from the children and teachers that they would prefer resources printed in colour, we also observed that the children clearly enjoyed colouring the prototype line drawings printed without colour. Another problem was that while we had hoped the children would take the books home and share what they were learning with their families, the teachers were worried about the books getting lost and the children not having them in class when they were needed. Our solution to both these problems was to create separate exercise books and textbooks. The final version of the children's (text) book was in colour, could be kept at school, and could be re-used by other classes. The exercise book (containing key learning goals for each lesson, exercises and a glossary) was printed in black-and white that could be coloured by the children, and could be taken home.

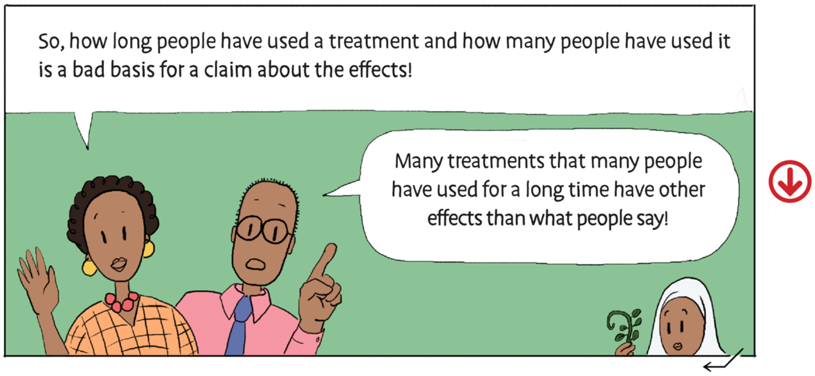
Identification

Initially we received many comments from the children in Uganda about the drawings, particularly about John and Julie, with whom they did not identify. However, with subsequent iterations of the children's book, the children identified with John and Julie. Similarly, both the teachers and children expressed that the resources felt like they were appropriate for them, increasingly with each iteration.

We were uncertain to what extent children at the international school in Norway would find the characters and the story, which was set in an East African context, relevant to them. To our surprise, we found that some of the children when asked where they thought the setting for the story was did not notice that it was in Africa. Others we spoke to were pleased that the story was set in Africa rather than in North America or Europe, which was the setting for most of the books they used.

Conclusions

Our findings suggest that with the iterative revisions of the IHC primary school resources, users - both children and teachers - experienced the resources as useful, easy to use, understandable, credible, desirable, and well suited for them. We believe there are two closely related reasons why we could achieve this. First, our grant application did not include a specification of what we were going to create. Instead, we described our goals and the methods that we would use to develop resources. This allowed us ample time (two years) to generate and prototype ideas and then to iteratively design, pilot and user-test, analyse, and redesign these resources.



Extra example

Ronald's claim: "Drinking fish oil keeps you healthy! I am sure because lots of people, for many years, have drunk fish oil to stay healthy!"
Treatment: Drinking fish oil
Effect: Having better health
Basis for Ronald's claim: How long people have used fish oil and how many people have used it
Explanation: Ronald's basis for his claim is bad, so his claim is unreliable. It is possible that fish oil does not make your health better, even though many people have used it for many years.

Lesson 3: Other bad bases for claims about treatments (Part 1) 53

Instructions and notes for teachers

Background: Health researchers have found that some herbs have good effects. For example, cream of hot pepper reduces back pain. However, they have found that other herbs have little or no good effects, and some have bad effects. For example, chewing betel nut causes cancer in the mouth. There are many herbs that they have not studied carefully.

Page **97** | Teachers' Guide

Fig. 7 Background section of each chapter for teachers

Second, we used a user-centred design approach with a multidisciplinary team and engagement of users throughout the development process. The research team included health service researchers with diverse backgrounds, designers, and a journalist. We collaborated closely with a teachers' network, a journalists' network [39], policymakers, and education researchers. We also piloted and user-tested the resources in schools in four countries. This broad range of feedback helped us create resources that increasingly resonated with these diverse communities.

The main limitation to the study was time constraint, in terms of tying the design schedule to the already busy school schedule. This also meant that only schools that were willing to avail time and participated in the development may not have been representative of the larger population. However in a follow up study, we have evaluated the effects of using the resources in a large randomised trial in Uganda (with 120 schools), that showed the intervention led to large improvements in the ability of both the children and their teachers to assess claims about treatments [36]. This trial excluded schools that participated in the development. Therefore, we can be fairly sure that input we gathered from participating schools was also representative for larger groups.

Supplementary information

Supplementary information accompanies this paper at https://doi.org/10.1186/s40814-020-00565-6.

Additional file 1. Overview of the development
Additional file 2. Structured Observation Form- Chapter one.
Additional file 3. Prioritisation of Key Concepts.
Additional file 4. Prototyping.
Additional file 5. User Experiences of the IHC resources.

Abbreviations

CARL: Critical thinking and Appraisal Resource Library; IB: International Baccalaureate Programme; IHC: Informed Health Choices Project

Acknowledgements

We are grateful for support for this research from the Global Health and Vaccination Research (GLOBVAC) programme of the Research Council of Norway, the English National Institute for Health Research and the James Lind Initiative for supporting IC and the African Academy of Sciences DELTAS Africa Initiative grant # DEL-15-011 to THRiVE-2, with funding from the Wellcome Trust grant # 107742/Z/15/Z for supporting DS and NKS. Miriam Grønli was responsible for colouring the final version of the children's resources and Nora Rosenbaum assisted. We are grateful Aisha Hashi, Sara Jaber, Rida Shah, Katie Tveiten, Cynthia Kateme, Mariam Nakato and Ann Babirye who helped test prototypes. Laetiticia Nyirazinyoye and Jenny Moberg helped with piloting and user-testing an earlier version of the learning resources. We want to thank Margaret Nabatanzi and Martin Mutyaba for their help with data collection. We would also like to thank the Informed Health Choices National Advisory Panel and the Teachers' Network. We are especially grateful to the teachers and children in Uganda, Kenya, Rwanda and Norway who helped with the development, user-testing and piloting of the Informed Health Choices primary school resources and especially the head teachers, the grade five children and teachers, at the following schools who participated in this: Bright Parents' Junior School, Kitante Primary School, Kisaasi Primary School, Nakivubo Primary School and Wampewo Primary School.

Authors' contributions

AN, DS, SR and ADO drafted this article manuscript. All the authors reviewed this article manuscript, provided input, and agreed on this final version. AN and DS were responsible for data collection in Uganda, MK led data collection in Kenya, MM led data collection in Rwanda and MO led data collection in Norway. All the investigators participated in the analyses of findings for each iterative cycle and agreed on consequences of the findings for the next version of the resources. MO drafted all resource texts; ADO and MO were responsible for final text decisions. SR illustrated, designed, and made final design decisions for the children's book, reminder poster and activity cards; AM designed and made final design decisions for the Teachers' guide. All authors read and approved the final manuscript.

Funding

The Research Council of Norway, Project number 220603/H10, funded the study. The Norwegian Institute of Public Health, recipient of the grant from the Research Council of Norway, is the coordinating centre for the Informed Health Choices project.

Availability of data and materials

All data will be available on reasonable request. (Extra data can be accessed on https://www.informedhealthchoices.org/learning-resources/).

Ethics approval and consent to participate

Teachers who were invited to participate in the pilot testing and user-testing were informed of the purpose of their participation before written permission was obtained. The children were given information about the project to take home for their parents and written permission was obtained for children selected to participate in the user-testing. Consent for the children to participate in the pilot testing was given by the head teachers and teachers. Children and their parents had the same right to refuse participation in piloting the IHC learning resources as they do for the use of any other learning resources used in the schools.
The study was approved by Makerere University Institutional Review Board and the Uganda National Council of Science and Technology as part of the *Supporting Informed Healthcare Choices in Low-income Countries Project (Grant no. ES498037)* in August 2013. Ethical approval was sought by the IHC project representatives in each of the other countries in compliance with national requirements.

Consent for publication

Written informed consent for publication of data and images was obtained from participants and from parents or guardians of the (children under 16) included in the study.

Competing interests

The authors declare that they have no competing interests.

Author details

[1]College of Health Sciences, Makerere University, Kampala, Uganda. [2]University of Oslo, Oslo, Norway. [3]Centre for Informed Health Choices, Norwegian Institute of Public Health, Postboks 222 Skøyen, 0213 Oslo, Norway. [4]Infodesignlab, Oslo, Norway. [5]Great Lakes University of Kisumu, Kisumu, Kenya. [6]University of Rwanda, Kigali, Rwanda. [7]Health Systems Research Unit, South African Medical Research Council, Cape Town, South Africa.

Received: 28 May 2019 Accepted: 31 January 2020
Published online: 10 February 2020

References

1. Sillience E, Briggs P, Harris PR, Fishwick L. How do patients evaluate and make use of online health information? Soc Sci Med. 2007;64(9):1853–62.
2. Irwig L, Irwig J, Trevena L, Sweet M. Smart health choices: making sense of health advice. London: Hammersmith; 2008.
3. Lloyd AJ. The extent of patients' understanding of the risk of treatments. Qual Health Care. 2001;10(1):14–8.
4. Weinstein N. What does it mean to understand a risk? Evaluating risk comprehension. J Natl Cancer Inst Monogr. 1999;25:15–20.
5. Robinson EJ, Kerr CE, Stevens AJ, Lilford RJ, Braunholtz DA, Edwards SJ, et al. Lay Public's understanding of equipoise and randomisation in randomised controlled trials. Health Technol Assess. 2005;9(8):1–192.
6. Austvoll-Dahlgren A, Nsangi A, Semakula D. Interventions and assessment tools addressing key concepts people need to know to appraise claims about treatment effects: a systematic mapping review. Syst Rev. 2016;5:215.
7. Sandoval WA, Sodian B, Koerber S, Wong J. Developing children's early compe-tencies to engage with science. Educ Psychol. 2014;49:139–52 Educ Psychol 2014; 49: 139–52. 2014.
8. Ruddock G, Sainsbury M. Comparison of the core primary curriculum in England to those of other high performing countries. Research Report DCSF-RW048. London: Department for Children, Schools and Families, 2008. Research report DCSF-RW048. London: Department for Children, Schools and Families, 2008; 2008.
9. Tamusuza A. Leaving School Early, The Quest for Universal Primary Education in Uganda. J Statisque Afr. 2011;11:110–51.
10. Burke K, Beegle K. Why children Aren't attending schools, the case of northwestern Tanzania. J Afr Econ. 2004;13(2):333–55.
11. Deininger K. Does the cost of schooling affect enrollment by the poor? Universal primary education in Uganda. Econ Educ Rev J. 2003; 22(3):291–305.
12. Evans DK, Popova A. What really works to improve learning in developing countries? An analysis of divergent findings in systematic reviews. . Policy Research Working Paper. Washington D. C: World Bank; 2015. p. 7203.
13. Nordheim L, Gundersen MW, Espehaug B, Guttersrud O, Flottorp S. Effects of school based educational interventions for enhancing adolescents' abilities in critical appraisal of health claims. a systematic review. PLoS One. 2016;11:0161485.
14. Semakula D, Nsangi A, Oxman M, Rosenbaum SE, Oxman AD, Austvoll-Dahlgren A, et al. Development of mass media resources to improve the ability of parents of primary school children in Uganda to assess the trustworthiness of claims about the benefits and harms of treatments. BMC Pilot Feasibility Stud. 2019;5:155.
15. Austvoll-Dahlgren A, Oxman AD, Chalmers I, Nsangi A, Glenton C, Lewin S, et al. Key concepts that people need to understand to assess claims about treatment effects. J Evid Based Med. 2015;8:112–25.
16. Nsangi A, Semakula D, Oxman AD, Austvoll-Dahlgren A, Oxman M, Rosenbaum S, et al. Effects of the informed health choices primary school intervention on the ability of children in Uganda to assess the reliability of claims about treatment effects: a cluster-randomised controlled trial. Lancet. 2017;17:31226.
17. Semakula D, Nsangi A, Oxman AD, Oxman M, Austvoll-Dahlgren A, Rosenbaum S, et al. Effects of the informed health choices podcast on the ability of parents of primary school children in Uganda to assess claims about treatment effects: a randomised controlled trial. Lancet. 2017;17:31225–4.

18. Nsangi A, Semakula D, Glenton C, Lewin S, Oxman AD, Oxman M, et al. Informed health choices intervention to teach primary school children in low-income countries to assess claims about treatment effects: process evaluation. BMJ Open. 2019;9(9):e030787.
19. Rosenbaum S. Improving the user experience of evidence: a design approach to evidence-informed health care. Arkitektur- og designhøgskolen i Oslo: Oslo; 2010.
20. IDEO. Human Centered Design Toolkit. 1st ed. San Francisco: IDEO; 2011.
21. Abras C, Maloney-Krichmar D, Preece J. User-Centered Design. In: Bainbridge W, editor. Encyclopedia of Human-Computer Interaction. Thousand Oaks: Sage Publi-cations, 2004; 2004.
22. Sanders E, Stappers P. Co-creation and the new landscapes of design. Co-design. 2008;2008(4):5–18.
23. Nsangi A, Semakula D, Oxman DA, Sewankambo KN. Teaching children in low income countries to assess claims about treatment effects; A prioritisation of key concepts. J Evid Based Med. 2015;8(4):173–80.
24. de Bono E. Lateral thinking. A textbook of Creativity. The Penguin Group; 1970.
25. Hillstrom K, Hillstrom LC. Brainstorming. In: Encyclopedia of small business, vol. 1. Farmington Hills: Gale Cengage; 2002.
26. Sauro J. A brief history of usability. Denver: MeasuringU; 2013.
27. Morville P. User Experience Design: honey comb model. Sematic Studios LLC 2004 (Updated June 21, 2004).
28. Chalmers I. Explaining the unbiased creation of treament comparison groups. Lancet. 2009;374:1671.
29. Krause L, Schwarts L, Woloshin S, Oxman A. Interactive tools to teach the public to be critical consumers of information about health care: What's out there? Abstracts of the 19th Cochrane colloquium, Madrid, Spain 157; 2011.
30. Castle JC, Chalmers I, Atkinson P, et al. Establishing a library of resources to help people understand key concepts in assessing treatment claims- The "Critical thinking and Appraisal Resource Library" (CARL). Plos One. 2017;12:e0178666.
31. Perrier MJ, Martin Ginis KA. Changing health-promoting behaviours through narrative interventions; a systematic review. J Health Psychol. 2018;23(11):1499–517. https://doi.org/10.1177/1359105316656243.
32. Houts P, Doak C, Doak L, Loscalzo M. The role of pictures in improving health communication. A review of research on attention, comprehension, recall and adherence. Patient Educ Couns. 2006;62:173–90.
33. Abrami PC, Bernard RM, Borokhovski E, Waddinton DI, Wade CA, Persson T. Strategies for teaching students to think critically; a meta analysis. Rev Edu Res. 2015;85:275–314.
34. McEwan PJ. Improving learning in primary schools of developing countries: a meta analysis of randomised experiments. Rev Edu Res. 2015;85:385–94.
35. Snilstveit B, Stevenson J, Phillips D. al E. interventions for improving learning outcomes and access to education in low and middle-income countries: a systematic review. London: International Initiative for Impact Evaluation; 2015.
36. Nsangi A, Semakula D, Oxman AD, Austvoll-Dahlgren A, Oxman M, Rosenbaum S, et al. Effects of the Informed Health Choices primary school intervention on the ability of children in Uganda to assess the reliability of claims about treatment effects; A cluster randomised trial. Lancet. 2017; 390(10092):374–88 Published online May 21.
37. Harden RM, Stamper N. What is a spiral curriculum? Med Teach. 1999; 21:141–3.
38. Bandura A. Self-efficacy: toward a unifying theory of behavioral change. Psychol Rev. 1977;84:191–215.
39. Semakula D, Nsangi A, Oxman AD, Sewankambo NK. Priority setting for resources to improve the understanding of information about claims of treatment effects in the mass media. J Evid Based Med. 2015;8(2):84–90.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Additional file 1. Overview of development

Methods and dates	*Participants*	*Description of activities*
PRIORITISATION OF KEY CONCEPTS		
Prioritsation of Key concepts workshop Uganda August 2013	The teachers' network in Uganda (24 teachers) and the research team (AN, DS, NS)	Members of the teachers' network attended a three day meeting where concepts were presented, discussed and prioritised using a pre-set criteria (19).
IDEA GENERATION AND EXPLORATORY PROTOTYPES		
Review of existing resources February 2013 to September 2014	The research team (AA, AM, AN, AO, CG, DS, SL, SR)	We searched for and reviewed existing resources (28-30).
Idea generation workshop *(Participatory collaboration)* February 2013	Researchers, teachers, and journalists from Indonesia, Nepal, Norway, Uganda, and the United Kingdom	At the 3-day kick-off meeting for the project, the research team together with invited teachers and journalists (18 people) discussed which concepts to focus on and brainstormed about potential resources. Journalists were participating as part of a parallel project to develop mass media resources.
Pilot testing in Norway *(Participatory observation)* February 2013	Approximately 30 ten-year old children and 2 teachers at an international school in Norway, and the research team (AA, AM, AN, AO, CG,JM, SP, SR, TT)	We developed materials and pilot tested an exploratory prototype (an experiment with colored candies)(28).
Prototyping workshop *(Facilitation & non-participatory observation)* September 2013	The teachers' network in Uganda (24 teachers) and the research team (AA, AM, AN, AO, CG, DS, NS, SR.)	This was a full-day workshop at which teachers brainstormed and created prototypes(19).
Pilot testing in Uganda *(Participatory observation)* September 2013	46 children and two teachers at a private school and 129 children and two teachers at a government school and the research team (AA, AM, AN, AO, CG, DS, MK, SR.)	We pilot tested the same exploratory prototype that was tested at the international school in Norway.
Meeting with the Ugandan National Advisory Board December 2013	15 members of the National Advisory Board for the project	This was a half day meeting during which ideas and barriers and facilitators to implementing these were discussed in the Ugandan context.

Methods and dates	*Participants*	*Description of activities*
Analysis of findings and idea generation December 2013 to January 2014	The research team (AA, AM, AN, AO, CG, DS, SL, SR)	We reviewed ideas that had been generated and their pros and cons. Based on this, we coded the ideas as 'good', 'not sure', or 'drop', and we identified principles to guide development of the resources.
Pilot testing in Uganda and Norway *(Participatory observation)* April 2014	27 nine to 13-year-old children and two teachers at 1 private school in Uganda and approximately 30 10-year-old children and 4 teachers at an international school in Norway, and the research team (AA, AO, AN, DS, MO, SR)	We developed and piloted a game designed to teach what a "testable question" is, using charts with relevant examples. The game included instructions, score sheets, timers, and a question bank. We also tested an activity where the children designed and carried out an experiment using paper airplanes.
Prototyping and pilot testing in Norway *(Participatory observation)* May to September 2014	Four 12-year-old girls, and the research team (AA, AM, AO, CG, SL, SR)	We prototyped a series of eight games and piloted these. We first tested each game by playing it ourselves. At each meeting with the children we introduced the relevant Key Concepts using a PowerPoint presentation, then played the game, then collected feedback.
Meeting and discussions with the members of the teachers' network September 2014	24 members of the teachers' network and 2 teachers who participated in the piloting of the materials, and the research team (AN, DS)	We updated the teachers about the prototypes we had user-tested and piloted. We asked them to try out the games in small groups of three before giving us feedback on what they thought about the progress being made so far. We also asked two teachers who had participated in the piloting of the materials to share their experiences with the rest of the group. The teachers discussed the challenges they faced and how they handled them.
Analysis of findings and idea generation September 2014	The research team (AA, AM, AN, AO, CG, DS, MO, SL, SR)	We reviewed our experience from prototyping and piloting the series of games. Based on this, we decided to develop a children's book to introduce the Key Concepts using a comic story and a teachers' guide.
Development of a partial prototype September to October 2014	The research team (AA, AM, AN, AO, CG, DS, MO, SL, SR)	We developed two chapters of the children's book and teachers' guide, including activities MO prepared a manuscript for each chapter, which was converted to a comic book by SR with exercises and activities. MO prepared a draft of each chapter of an accompanying teachers' guide and AM designed the guide. Each chapter was reviewed by the rest of the research team and went through three iterations.

Methods and dates	*Participants*	*Description of activities*
Pilot testing *(non-participatory observation)* **and user-testing in Uganda** October 2014	73 year-5 children and two teachers at a government school and 28 year-5 children and one teacher at a private school piloted the prototype, and the research team (AM, AN, AO, DS, MO, SR)	We pilot tested the first chapter of the children's book and teachers' guide, including activities, at both schools and the second chapter at the private school. We interviewed four 10 to 15-year-old children (two from each school) and all three teachers.
Analysis and idea generation November 2014	The research team (AA, AM, AN, AO, CG, DS, MO, SL, SR)	We coded findings from the pilot and user testing as "show stoppers" (very important problems), important problems, minor problems, positive feedback, or specific suggestions.
Feedback gathering meetings December 2014 to January 2015	National Advisory Board for the project and the teachers' network	AN, DS, and NS presented plans for a complete prototype of the IHC primary school resources and sought feedback and input. We conducted a half-day meeting with the National Advisory Board and a full-day meeting with the teachers. AN, DS and NKS recorded the sessions as the policymakers and stakeholders discussed.
VERSION 1 OF THE IHC PRIMARY SCHOOL RESOURCES		
Development of a complete prototype December 2014 to April 2015	The research team (AA, AM, AN, AO, CG, DS, MO, SL, SR)	We outlined a children's book with 10 chapters that would cover 24 Key Concepts. MO prepared a storyboard for each chapter, which was converted to a comic book by SR with exercises and activities. MO prepared a draft of each chapter of an accompanying teachers' guide and AM designed the guide. Each chapter was reviewed by the research team and went through three iterations.
Pilot testing *(non-participatory observation)* **and user-testing in Uganda** January to May 2015	67 year-5 e children and one teacher at a government school, and 32 year-5 children and one teacher at a private school, and the research team (AN, DS, research assistents)	We gave the teachers a chapter a week before they taught each lesson. The teachers who taught the lessons were not given any instruction other than what was in the teachers' guide. AN and DS observed each lesson. They recorded their observations using a semi-structured guide. They interviewed three children from each school and both teachers after each chapter was pilot tested using a semi-structured interview guide. There was an observer who took notes at each interview. All interviews were audio recorded.

Methods and dates	*Participants*	*Description of activities*
Feedback gathering meeting with the teachers' network May 2015	24 members of the teachers' network, 2 teachers that had participated in the piloting of the materials, and 2 teachers that had participated in the user testing of the materials	This was a one day meeting where we grouped the teachers in small groups of about three teachers per group and asked them to read a chapter and give us feedback on what needed to be improved and what should be dropped. AN, DS, and NKS recorded the sessions as teachers gave their feeback.
Analysis and idea generation May 2015	The research team (AA, AM, AN, AO, CG, DS, IC, MK, MO, NS, SL, SR)	AN and DS entered the findings in a spreadsheet. For each finding, AN, AO, DS, MO, and SR coded its importance (very important, important, or less important); whether it was a problem, an idea, or positive feedback; and whether it applied to the entire book, a specific chapter, or was a repeat of a previous finding. The findings were summarized for the research team and the major findings and plans for the second version were discussed and agreed.
VERSION 2 OF THE IHC PRIMARY SCHOOL RESOURCES		
Development of the second complete prototype June to August 2015	The research team (AA, AM, AN, AO, CG, DS, IC, MK, MO, NS, SL, SR)	MO prepared a revised draft for each chapter, SR revised the drawings and AM revised the design of the guide. Each chapter was reviewed by the research team and went through three iterations.
Pilot testing *(non-participatory observation)* **and user-testing in Uganda** September to December 2015	96 year-5 children and one teacher at a government school and 109 children and one teacher at another government school, and the research team (AN, DS, research assistents)	We gave the teachers the materials in two parts (chapters 1 to 4 and chapters 5 to 10). The teachers who taught the lessons were not given any instruction other than what was in the teachers' guide. AN and DS observed each lesson. They recorded their observations using a semi-structured guide. They reviewed the children's completed exercises from the previous lesson and they interviewed two children from each school and the two teachers after each chapter was pilot tested using a semi-structured interview guide. They also interviewed an additional teacher who did not teach the lessons. There was an observer who took notes at each interview. All interviews were audio recorded.

Methods and dates	***Participants***	***Description of activities***
Pilot testing *(non-participatory observation)* **and user-testing in Kenya** September to December 2015	30 children between 10 to 14 years old and one teacher, and the research team (MK and research assistents)	The teachers was initially provided with the teachers' guide prior to the pilot and user testing to enable him familiarize himself with the materials. The relevant lesson for each week was covered by the children under the guidance of the teacher. Following each lesson, between 4 and 6 pupils were interviewed using a semi-structured guide. An observer recorded the interviews with the children. A semi-structured guide was used to report the lesson findings. Some of the lessons and interviews with pupils were recorded.
Pilot testing *(non-participatory observation)* **and user-testing in Rwanda** September to December 2015	33 year-5 children (10 to 12-years old in year-5) and one teacher, and the research team (MM and AU)	MM and AMU used the same methods for the pilot study as described above for Uganda and Kenya. Two or three observed 10 class sessions (non participatory observation). Then after each lesson, one of them interviewed and the other taken notes for three of the children in a focus group and the teacher (individually) in an interview using a retrospective think aloud technique, going through each page of the book with a semi-structured interview guide. The interviews were audio recorded.
Pilot testing *(non-participatory observation)* **and user-testing in Norway** September to December 2015	Three year-7 classes with 15 to 18 children from many different countries in each class and two teachers (one who taught two different classes) at an English-language international school and the research team (MO, SR, AO, AA)	One or two researchers observed each lesson in each class using a structured data-collection form and then entered findings into a spreadsheet. We interviewed four children selected by one of the teachers with one person conducting the interview and one observer; we interviewed each teacher twice, and we collected verbal feedback from each class after they completed all 10 chapters; we reviewed their completed exercises, and we interviewed the school's head of science. All the interviews were semi-structured using interview guides and were recorded.
Update on current activities and feedback gathering meeting of teachers' network members December 2015	24 members of the teachers' network	At this full-day meeting, teachers were updated on the progress before being divided in groups of about three. We asked each group to look at the entire chapter of the teachers' guide assigned to their group and provide the research team (AN, DS, and NKS) with feedback on what needed to be addressed.

Methods and dates	*Participants*	*Description of activities*
Analysis and idea generation December 2015 to January 2016	The research team (AA, AM, AN, AO, AU, CG, DS, IC, MK, MM, MO, NS, SL, SR)	For each finding, AN, AO, DS, MM, MO, and SR coded its importance (very important, important, or less important); whether it was a problem, an idea, or positive feedback; and whether it applied to the entire book, a specific chapter, or was a repeat of a previous finding. The findings were summarized for the research team and the major findings and plans for the second version were discussed and agreed.
VERSION 3 OF THE IHC PRIMARY SCHOOL RESOURCES		
Development of the final set of learning resources January to March 2016	The research team (AA, AM, AN, AO, AU, CG, DS, IC, MK, MM, MO, NS, SL, SR)	MO prepared a storyboard for each chapter with exercises and activities. This was converted to a comic book by SR. MO prepared a draft of each chapter of an accompanying teachers' guide and AM designed the guide. Each chapter was reviewed by the research team and went through three iterations.

IHC

PILOT OBSERVATION FORM

SCHOOL RESOURCES

CHPT.1

SECTION A:	**Pre-lesson**
Observer:	AN
School:	(X) primary school
Scheduled start time of lesson:	2:00 pm
Scheduled end time of lesson:	2:40 pm
Number of children:	100
Age range of children: *From youngest to oldest child*	10 years (youngest) 15 years (oldest)
Number of benches: *This is so we know about how much space each child had.*	24 benches
Number of teachers in the room:	one
SECTION B:	**Start of lesson**
Actual start time of lesson:	2: 13 pm
What did the teacher do before the class started reading the chapter? For how long? *E.g. makes jokes, ask the children questions or give his own summary of the book or a chapter.*	- The teacher asked the children; to welcome the visitors - The teacher also asked the visitors to introduce themselves. - The teacher started by asking the children to name the body parts. - The teacher introduced the book by writing the title on the black board. - The teacher started introducing the words “treatment” and Effects. Doing it in both Luganda and English i.e. Treatment is obujanjabi. Giving an example of how we treat malaria. The children named malaria treatment like Quartum and then discussed effects i.e. ekivamu in luganda. The children named some effects like sweating, vomiting etc. The teacher also named some good effects for example getting better.
SECTION C:	**Reading**

	Page 5; the children struggled with the word injuries on page 5 in the children's book. "Injuries" – Children asked the teacher. Teacher replied in Luganda "Injury is okukosebwa"
Page 12--Chapter title page	The teacher skipped reading the title "Claims, Comparisons and Choices; A health science book".
Page 13--"Health"	The children were able to read the word health and the teacher compared the two pictures on the page; i.e. "One child is healthy and doing back summer", and the other child is unwell.
Page 14 to 15--"Treatments"	- The children were able to read the word treatment. - They struggled with reading the word "crutches". Teacher had to explain what crutches meant in Luganda and what it is used for. - The word vegetable; some children struggled with that word
Page 16 to 19--"Effects"	Effects; Teacher asked the children to first name treatments that they have read in the book; the children mentioned; using an equipment The children started reading again; they read effects well and then continued reading the book. The children named the good and bad out of playing football; i.e. when you play football…(children were asked to name them) Good Bad -Being strong -Getting a fracture
Page 20 to 21--"things people say"	The children were able to read the word herbs without difficulty. Teacher kept referring to both words; i.e. treatment is obujanjabi
Page 22 to 24--"This book"	The children kept reading while the teacher would stop and interpret difficult words into Luganda for the children.
SECTION D:	**Activity**
	Teacher asked children;

	To discuss the examples of treatment; i.e. children mentioned exercising, eating fruit and vegetables, getting an operation, dressing a wound; using plasters on a broken arm.
SECTION E:	**Exercises**
	The children started doing their exercises in their books. Exercise 3; Children were a bit unsure; the teacher asked them to choose the most correct answer and write it down. The teacher kept explaining what was required and asked the children if they had understood the instructions to which the children replied yes they had.
SECTION F:	**Post-lesson**
Who read aloud?	All the children read aloud
Actual end time of class: *When did the teacher leave the children to work on their own or do other things?*	One hour; 3:00 pm
About how long did the class spend reading the story?	30 minutes
About how long did the class spend doing the activity?	10 minutes
About how long did the class spend doing exercises with the teacher in the room?	20 minutes
About how long did the children break for?	0 minutes
Did the children seem interested or disinterested in the lesson? How so?	The children were very excited, partly because they each got their own book.
How did the teacher engage the children?	The teacher encouraged the children especially the back benchers to answer questions as they arose.
How did the teacher use the blackboard?	The teacher kept writing key words on the blackboard.

What did the children do with the books at the end of the class? *Did they take them home or hand them in?*	The teacher collected the books and told the children that she would be marking them later.

Additional file 3. Prioritising Key Concepts

We started with the list of 32 Key Concepts identified at the beginning of this project (15), organised in six groups:

- Recognising the need for fair comparisons of treatments
- Judging whether a comparison of treatments is a fair comparison
- Understanding the role of chance
- Considering all the relevant fair comparisons
- Understanding the results of fair comparisons of treatments
- Judging whether fair comparisons of treatments are relevant

To determine how many of these Key Concepts the prototypes should teach, we first consulted with the teachers' network at a workshop in August 2013. They found all six groups of concepts to be relevant for year-5 (10 to 12-year-old) children (19). Based on input from the teachers, we judged that primary school children could learn 24 of the 32 concepts. Members of the research team made these judgements in a face-to-face meeting in September 2013, using informal discussion to reach a consensus.

However, 24 concepts proved to be too much to learn in a school term. The early prototypes we created had too many concepts per lesson and took too long to teach in a normal school hour (40 minutes). We also observed that the teachers needed time to repeat material from previous lessons. After each round of proto typing, we eliminated more Key Concepts from our list.

We decided which ones to eliminate by considering the importance of the concepts and the difficulty that the children had learning them. The importance of the concepts was based on judgements made by members of the research team by:

- Each person individually identifying which of the 24 Key Concepts they considered most important
- Compilation and discussion of those judgements
- Voting on the concepts
- Reaching a consensus by informal discussion

At a face-to-face meeting in May 2015, we reached agreement that eight of the concepts were most important for our target population in Uganda. Three members of the research team also reviewed data from our piloting and user-testing and identified concepts that appeared to be too difficult to teach to 10 to 12-year-old children. Later, based on feedback from piloting the resources, we considered how the concepts were grouped in the lessons and the number of concepts being taught in each lesson.

We ended up using 12 of the 24 concepts in the final version of the resources (Box 1) and reorganised them into three groups to simplify and clarify their purpose:

- **CLAIMS**: "questions you should ask when someone says something about a treatment"
- **COMPARISONS**: "questions that health researchers ask to find out more about the effects of treatments"
- **CHOICES**: "questions that you should ask when you are choosing whether to use a treatment"

Additional file 4. Prototyping and pilot testing in Uganda and Norway

Prototyping workshop with teachers' network in Uganda

All 24 members of the teachers' network attended this full-day workshop, facilitated by eight members of the IHC team (Figure 1). We guided the teachers through a structured process that led them to create a persona, identify barriers and facilitators, generate ideas about resources, and rapidly prototype some of the ideas. Persona are fictional characters created to represent a user type, often used in digital design development (31).

Figure 1. Prototyping workshop with teachers' network in Uganda

Some of the categories of ideas teachers came up with were: assemblies, community involvement, family involvement, debates, use of drama, games, experiments, reading materials, video, teacher education, and curriculum development.

Although the resulting prototypes did not represent resource ideas we could use directly (as the ideas depended too much on teacher's in-depth prior knowledge gained through their participation in our workshop), we developed a better understanding of teachers' perspective on our work as well as many insights into the school and community setting in Uganda. Key messages were that:

- There was a paucity of available materials in schools, including paper and printers.
- Songs, drama, and storytelling were popular methods for conveying health messages.
- Children might be expected to share their knowledge with their family or community.

We found that it was difficult for some people to grasp that we were focusing on teaching health literacy (specifically critical thinking about treatment claims) and not health promotion (teaching about the benefits of specific "treatments" such as handwashing or using insecticide treated bed nets).

Drawing on what we learned, we focused our next efforts on creating prototypes of highly structured materials that we could produce and provide cheaply for each school, and that did not rely on teachers' prior knowledge about the Key Concepts. We continued to try to create games.

Prototyping and pilot testing in Uganda and Norway

We developed and piloted a game designed to teach what a "testable question" is and tested an outdoor activity where the children designed and carried out an experiment using paper airplanes. We piloted and user-tested these games at a Ugandan school and at an international school in Norway (Figure 2).

Figure 2. Pilot testing the prototype of a "testable question" game in Norway

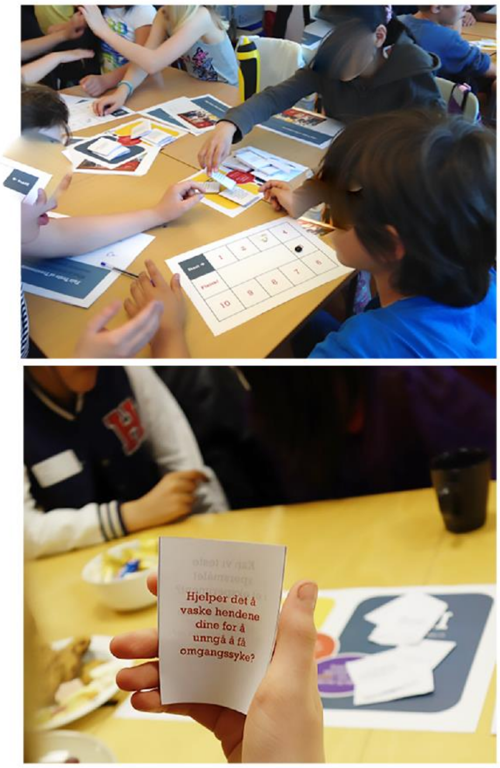
Level 1 – Asking testable questions
Make a set of cards with questions that are testable and ones that are not. Take turns picking cards and saying whether the question is testable or not.

Round 1 – Start with a prepared set of questions with correct answers and a teachers' guide.

Round 2 – Kids make up their own questions + correct answers (+ a teachers' guide).

- Scoring option 1 (competitive): 1 point per correct answer and -1 point for wrong answers. Person or team with the most points wins.

- Scoring option 2 (collaborative): 1 point per correct answer and -1 point for wrong answers. Class/group needs to get a fixed number of points before moving on to the next level.



Then we developed eight prototypes for simple games with increasing difficulty, designed to teach one concept each. These were based on materials that could be printed on paper. We tested these out in a series of meetings over two months with a group of four 12-year-old girls in Norway. At each meeting with the children we introduced the relevant Key Concepts using a PowerPoint presentation, then played the game, then collected feedback.

Development of a partial prototype of a comic book and teachers' guide

We developed two chapters of a children's comic book and teachers' guide, with an activity description in each and a separate "activity booklet". We carried out pilots (with non-participatory observation) and user-tests in two Ugandan schools as described in Additional file 1, followed by data analysis (Figure 6).

Figure 3. Pilot testing a partial prototype of comic book and teacher's guide

Negative user experiences
We categorized several findings as very important negative findings ("show stoppers"). For example, in one of the schools, the class spent three and a half hours on the first chapter alone, with the teacher often improvising off-script and coming up with examples that led the children to ask many unrelated and misleading questions; for example:

> *"Does using the pit latrine while pregnant cause some mothers to lose their children, true or false?"*

At the second school, the outdoor activity became very chaotic and took too long, leaving no time for discussion about what the children had learned. Also, some of the key terminology (e.g. 'treatment' and 'outcome') was too complicated for the children.

Other important negative findings included: The size of the text was too small in both books. Some children were reading a comic book for the first time and had difficulty following the flow of text and understanding the comic layout. We observed that some of the examples we used misled the children, for example:

> *"[I've learned] not to eat so much if you want to run fast".*

Other negative findings included some of the words we had used that were unfamiliar to both teachers and children (e.g. 'Prattle' and 'gobbling'). Typical comic book sound words like 'bling' and 'whoosh' were also unfamiliar. In addition, some of the children did not want to identify with some of the characters, because they were not like them; for example:

> *"The girl has not combed her hair". "John's shirt is not tucked in".*

When asked about the usability of the product, the teacher said:

> *"First experience was difficult, will be better next time."*

Positive user experiences

A very important positive finding was the fact that both teachers and children were enthusiastic about the use of comics. They liked the pictures and the way the information was presented which broke the routine of blackboard teaching.

> *"We didn't just talk and use the chalkboard, there was a book for each child."*

When we asked teachers about the desirability of the product and its suitability for their pupils, there were several positive findings:

> *"The way the children were thinking, it was extraordinary."*
> *"The level is good for these children."*

However, it became clear to us from classroom observations and user-testing interviews, that the children had obviously not understood some of the content in the book; for example:

> *"The question I have about what John and Julie learn in this chapter is: Why did John and Julie use a coin to divide ten and a hundred people?" Why didn't they use counting to divide the people?"*

Many children coloured in their books. The use of animals (e.g. "Prattle", the parrot) seemed to be interesting to them, but there was disagreement among teachers and children about the appropriateness of including animals that could talk in the story. It was "not real" as stated by some teachers:

> *"What is the talking bird doing here?"*
> *"The children will not take this seriously because of this bird here, they will think it is a joke."*

The children, however, liked Prattle; as one stated during a user-testing interview:

> *"I think that bird is so funny because it is saying many funny claims."*

Suggestions from teachers categorized as very important constructive findings included: providing extra examples, adding numbers to the frames so that the children know which direction to read, having a thicker paper on the cover of the teachers' guide, colouring the pictures, and breaking the content into smaller teachable units.

Additional file 5. User experiences of the IHC primary school resources

Version 1 of the IHC primary school resources

Negative user experiences

The objective of the lessons and materials was misunderstood, as we found out in our interviews with some of the children. One of the very important negative findings was an incorrect understanding of what the book was about. Some children expected they were going to evaluate treatments themselves or learn about practicing medicine.

> *"I am going to learn being a scientist and being a health treater and treating other people"*
>
> *"I think I will learn more about science and health. The needs of being a doctor and how I will become one".*

Two chapters were too long to be completed in one school hour (lesson), and a teacher said there were too many lesson goals. There was much concept-related vocabulary that some children did not understand despite the definitions and translations; e.g. claim, outcome, substitute outcome, assumption, unreliable, careful summary. Additionally, some children struggled with simpler English words; e.g.: expert, also, normally, reward. Feedback suggested that English skills would likely be even poorer in more rurally located schools. One child stated during an interview:

> *"This book is for a school like ours which knows hard English, but not for the village schools".*

Another important negative finding in this version was the volume of the teachers' guide and the content in the children's book. Some teachers suggested that the chapters should be split into manageable units. One of the pilot teachers attending the network meeting stated:

> *"I found the information in the chapter to be too much when delivering the lesson, it was difficult to have the children read the chapter and also do all the exercises in the 45 minutes of the lesson."*

A member of the teachers' network said:

> *"This lesson plan is difficult to follow."*

Other important negative findings for this version included the use of the English language, which the target audience (teachers and children) still found difficult to understand. One of the children said in an interview:

> *"The difficult word was ingredients. I don't know what it means."*

In addition, teachers were not comfortable with some of the English words used in the teachers' guide and in the children's book. One stated:

> *"The children's book had some confusing and difficult words like 'effects'. Yet, in my own understanding, I had a different meaning of the word effect."*

Other negative findings included that the illustrations (which were still just sketches drawn with pencil) were unclear and the hand written text was difficult for many to read. Some of the children were unsure about how to fill in the answers to the exercises and others struggled to read the comic in the right direction, despite the simplified layout.

In addition, we also observed teachers improvising from the text and offering incorrect examples and analogies. Teachers expressed wanting step-by-step instructions that could also be followed by a stand-in teacher who was unfamiliar with the project and more in-depth information to help them be prepared if challenged by children.
One teacher suggested making the professor characters in the story doctors, and others suggested changing their names.

Positive user experiences
The teachers felt this version was a big improvement over the partial prototype, and that the content became clearer from chapter to chapter.
Using a combination of pictures and text seemed to be helpful both from the teachers' and children's perspective. One child stated:

> *"The words help you understand the pictures."*

Another important positive finding was that the exercises seemed to be the right level of difficulty. Both teachers and children valued them and suggested that more exercises should be included. One of the teachers said this about the exercises in the book:

> *"Good, perfect, children will be able to do this and it even helps them improve on the area of English."*

The exercises in the book seemed tuned to the children's literacy levels with many children correctly completing the exercises within 5-10 minutes. The Luganda translation was also perceived as helpful. One child said:

> *"Effect - it means Ekivaamu. It is good to tell us the meaning of the word."*

Children loved the parrot. One of them suggested:

> *"More pictures about birds."*

Version 2 of the IHC primary school resources

Version 2 - Negative findings
Some of the very important negative findings for this version included: feedback from the pilot teachers who felt that some chapters in this version were very long, particularly chapter 8. One pilot teacher stated:

> *"Some of these chapters, like chapter 8, need to be divided into smaller chapters. They take a lot of time."*

A member of the teachers' network said:

> *"The content is too wide".*

Another very important negative finding arose from observations. In a class of about 109 children, a teacher took a lot of time organizing the children into groups to participate in the "CLAIM Game". When asked about this during the interview after the lesson, the teacher said:

> *"Organising children and preparing them for the game is difficult in a class like mine."*

An important negative finding observed during the lessons was that the use of the teachers' guide differed from teacher to teacher in the pilot schools. One of the teachers was observed not using the guide during the lesson. The teacher only glanced through it when needed and mostly used the children's book as the class read aloud. During our interview after the lesson, the teacher said:

> *"Having to use two books in a lesson is very hard. You have to look at the guide and also see what the children are reading. Why don't we make this into one book?"*

The language used in this version of the resources was still found to be a bit complicated for the children, as this was observed during the lessons when children were required to read aloud and struggled with some of the words they found difficult to read and to pronounce; for example, "surrogate outcome". During our consultations at meetings with members of the teachers' network, one said:

> *"Some of the chapters have very many new words; for example, chapter 2. Words like 'health research', 'claim', 'experience', 'reason', 'eh', those are so many for the children."*

During one of the pilot lessons, we observed that the teacher was hurriedly marking a few of the children's books as they were attempting the exercises during the lesson. This allowed no time for the children to read and understand, if they were to be among the few whose books would be marked. The teacher said during the interview:

> *"There are so many exercises for the children to do after each lesson, it makes it very difficult for me to mark all of them."*

We observed a negative finding when one of the pilot teachers struggled to organize a class of over 100 children to participate in the CLAIM game. The children struggled to understand the instructions as they were laid out. In the interview after the lesson, the teacher said:

> *"I am worried about what you thought when you were watching me. I tried to organize myself last night for this activity but it was very difficult. I had to mark the children's exercise books and also prepare for the activity. It is very difficult to do in this big class."*

Version 2 - Positive findings

A positive finding was that the exercises in the children's book were considered appropriate and engaging for year-5 children. One of the pilot teachers said:

> *"The exercises are very well suited for these children."*

Having new terms in three different languages in the textbook (English, Luganda, and Kiswahili) and explaining difficult words in a language that year-5 children could understand improved comprehension of the content being taught. A pilot teacher said:

> *"Using Luganda to explain the word 'effect' - 'ekivaamu' - made it simple for me to understand and then also explain to the learners."*

II

Effects of the Informed Health Choices primary school intervention on the ability of children in Uganda to assess the reliability of claims about treatment effects: a cluster-randomised controlled trial

Allen Nsangi, Daniel Semakula, Andrew D Oxman, Astrid Austvoll-Dahlgren, Matt Oxman, Sarah Rosenbaum, Angela Morelli, Claire Glenton, Simon Lewin, Margaret Kaseje, Iain Chalmers, Atle Fretheim, Yunpeng Ding, Nelson K Sewankambo

Lancet 2017; 390: 374–88

Published Online May 21, 2017 http://dx.doi.org/10.1016/S0140-6736(17)31226-6

See Comment page 336

College of Health Sciences, Makerere University, Kampala, Uganda (A Nsangi MSc, D Semakula MD, Prof N K Sewankambo MD); **Centre for Informed Health Choices, Norwegian Institute of Public Health, Oslo, Norway** (A D Oxman MD, A Austvoll-Dahlgren PhD, M Oxman BJ, S Rosenbaum PhD, A Morelli MA, C Glenton PhD, S Lewin PhD, Prof A Fretheim PhD, Y Ding PhD); **University of Oslo, Oslo, Norway** (A Nsangi, D Semakula, Prof A Fretheim); **Infodesignlab, Oslo, Norway** (A Morelli); **Health Systems Research Unit, South African Medical Research Council, Cape Town, South Africa** (S Lewin); **Great Lakes University of Kisumu, Kisumu, Kenya** (M Kaseje PhD); and **James Lind Initiative, Oxford, UK** (Sir I Chalmers DSc)

Correspondence to: Dr Andy D Oxman, Centre for Informed Health Choices, Norwegian Institute of Public Health, Oslo N-0403, Norway **oxman@online.no**

Summary

Background Claims about what improves or harms our health are ubiquitous. People need to be able to assess the reliability of these claims. We aimed to evaluate an intervention designed to teach primary school children to assess claims about the effects of treatments (ie, any action intended to maintain or improve health).

Methods In this cluster-randomised controlled trial, we included primary schools in the central region of Uganda that taught year-5 children (aged 10–12 years). We excluded international schools, special needs schools for children with auditory and visual impairments, schools that had participated in user-testing and piloting of the resources, infant and nursery schools, adult education schools, and schools that were difficult for us to access in terms of travel time. We randomly allocated a representative sample of eligible schools to either an intervention or control group. Intervention schools received the Informed Health Choices primary school resources (textbooks, exercise books, and a teachers' guide). Teachers attended a 2 day introductory workshop and gave nine 80 min lessons during one school term. The lessons addressed 12 concepts essential to assessing claims about treatment effects and making informed health choices. We did not intervene in the control schools. The primary outcome, measured at the end of the school term, was the mean score on a test with two multiple-choice questions for each of the 12 concepts and the proportion of children with passing scores on the same test. This trial is registered with the Pan African Clinical Trial Registry, number PACTR201606001679337.

Findings Between April 11, 2016, and June 8, 2016, 2960 schools were assessed for eligibility; 2029 were eligible, and a random sample of 170 were invited to recruitment meetings. After recruitment meetings, 120 eligible schools consented and were randomly assigned to either the intervention group (n=60, 76 teachers and 6383 children) or control group (n=60, 67 teachers and 4430 children). The mean score in the multiple-choice test for the intervention schools was 62·4% (SD 18·8) compared with 43·1% (15·2) for the control schools (adjusted mean difference 20·0%, 95% CI 17·3–22·7; p<0·00001). In the intervention schools, 3967 (69%) of 5753 children achieved a predetermined passing score (≥13 of 24 correct answers) compared with 1186 (27%) of 4430 children in the control schools (adjusted difference 50%, 95% CI 44–55). The intervention was effective for children with different levels of reading skills, but was more effective for children with better reading skills.

Interpretation The use of the Informed Health Choices primary school learning resources, after an introductory workshop for the teachers, led to a large improvement in the ability of children to assess claims about the effects of treatments. The results show that it is possible to teach primary school children to think critically in schools with large student to teacher ratios and few resources. Future studies should address how to scale up use of the resources, long-term effects, including effects on actual health choices, transferability to other countries, and how to build on this programme with additional primary and secondary school learning resources.

Funding Research Council of Norway.

Introduction

Good health depends partly on people making good choices. Good choices depend on health literacy—ie, people's ability to obtain, process, understand, and judge the reliability of relevant health information. However, people often lack the ability to judge the reliability of information about the effects of treatments, and they tend to overestimate treatment benefits and underestimate treatment harms.[1] Low health literacy is associated with poor health outcomes and poor use of health-care services.[2] Improving health literacy, and particularly people's ability to assess claims about treatment effects, has the potential to reduce unnecessary suffering and to save billions of dollars every year.[3–5]

Most health information offers instructions or claims without adequate information for people to make informed choices. Meanwhile, much health and science education, which could teach people to assess

Research in context

Evidence before this study

At the start of the project (June 21–22, 2013), we searched the Cochrane Library, MEDLINE (Ovid), and ERIC for any quantitative study that measured the ability of participants to assess claims about the effects of treatments. We also contacted key researchers working in related research areas. We did not include reports in languages other than English or the Scandinavian languages. We did not find any studies that evaluated a primary school intervention to teach children to critically appraise treatment claims or make informed health choices, in any country.

A systematic review (Abrami and colleagues, 2015) of the effects of strategies for the development and enhancement of critical thinking skills at any age and in any setting found 49 studies of such strategies for teaching children aged between 6 and 10 years. However, none of these strategies focused specifically on health literacy. An overview (Evans, 2015) of six systematic reviews of educational interventions in low-income and middle-income countries included 227 studies that reported learning results. None of these studies addressed health or science literacy, or critical thinking more broadly. Systematic reviews (Austvoll-Dahlgren, 2016; Nordheim, 2016) of teaching children critical appraisal skills in relation to health have not identified studies that evaluate the effects of strategies for teaching these skills to primary school children.

Added value of this study

This is the first randomised trial to evaluate any intervention to improve the ability of primary school children anywhere to assess claims about treatments. We found a large effect: an increase of nearly 50% in the proportion of children with a passing score on a test that measures their ability to assess treatment claims. This corresponds to an effect size that was well above the average for other critical thinking interventions for any type of student in any country. No adverse events were reported. As with any school activity, the time that is used for this intervention (13 h over a 12-week school term) must be taken away from other activities. The cost of the intervention (about US$4 per child) is substantial relative to current levels of expenditures per primary school child in Uganda and other low-income countries.

Implications of all the available evidence

It is uncertain what the long-term effects of using the Informed Health Choices primary school resources are, what if any effect the programme will have on actual health choices and outcomes, or how transferable the findings of this study are to other regions and countries. Additionally, although the cost of the intervention is small, it is a substantial cost compared with the cost of school in Uganda. Nonetheless, being able to think critically about treatment claims (and generally) has an intrinsic value. School authorities, teachers, and children in the study indicated that they consider it important. We recorded a large effect on critical thinking about treatment effects, which was the primary outcome. Future research should address how best to scale up use of the resources, their suitability and effects in other countries, and how to build on these resources with additional primary and secondary school resources.

health claims, tends towards rote learning rather than critical thinking.[6] Economically disadvantaged people suffer disproportionately if they are unable to make informed health choices, as they can least afford to waste resources.

Teaching primary school children how to assess claims about the effects of treatments might be an effective strategy for several reasons. First, children are capable of learning about fair tests (ie, controlled investigations) and critical appraisal.[7] Indeed, teaching these basic skills is already part of school curricula in some countries.[8] Second, by targeting primary school children, it is possible to reach a large segment of the population (before many leave the education system and become difficult to reach). Large numbers of children drop out after primary level in low-income countries.[9] Third, teaching children at primary school level to assess claims about treatments can capitalise on the time these children have available for learning. Conversely, young people and adults have increasing demands on their time and it becomes increasingly difficult to teach them to think critically about treatment claims if they lack a foundation. They have less time to learn and must learn more at once. Moreover, erroneous beliefs, attitudes, and behaviours developed during childhood might be resistant to change later, when children become adults.[10] Fourth, teaching critical thinking skills to young children improves their academic achievement, and these effects are larger for low-achieving children.[11,12] Finally, learning to think critically about treatment claims can prepare children to contribute to well informed health policies as citizens, as well as to make their own personal health choices.

Although primary school children are taught about fair tests and critical thinking in some countries,[8] the focus is not on health or assessing claims about the effects of treatments. The aim of this study was to test the effects of using learning resources on the ability of children to assess claims about the effects of treatments. In a separate paper,[13] we will report a process evaluation in which we investigate factors that might have influenced the effect of the intervention, ways of scaling up effective use of the resources, and other potential beneficial and harmful effects of using the resources.

Methods

Study design

In this two-group cluster-randomised trial, we included 120 primary schools in the central region of Uganda. Ethics approval was obtained from the School of Medicine's institutional review board at Makerere University College of Health Sciences and the Uganda National Council for Science and Technology. We obtained approval to do the trial from the Ugandan Ministry of Education before recruiting study participants.

For the **random number generator** see http://www.sealedenvelope.com

Participants

Primary schools in Uganda normally fall under a regional authority headed by a district education officer, who is the primary contact between the Ministry of Education and the schools in that region. For this study, we obtained an introductory letter from the Permanent Secretary at the Ministry of Education introducing us to the district education officers in the region. We informed the district education officers about the project and asked them to provide us with a list of all primary schools in the region. We used this list as our sampling frame to identify eligible schools.

We used a multistage sampling technique in which we first drew a random sample of four districts from all 24 districts in the region (appendix 1). In the second stage, we randomly sampled schools proportionately from lists of the selected districts, stratifying by school location (urban, semi-urban, or rural) and ownership (private or public). With the help of the district education officers we generated a list of 2029 eligible schools in those four districts. We excluded eight international schools, five special needs schools for children with auditory and visual impairments, four schools that had participated in user-testing and piloting of the resources, 160 infant and nursery schools, and one school for adult education. For practical reasons, we also excluded 753 schools that were difficult for us to access in terms of travel time. We then randomly selected 170 of the remaining schools.

See Online for appendix 1

We (AN and DS) visited schools that were selected for recruitment, taking with us a letter of introduction from the respective district education officer. We provided the head teacher of each school with information about the study and obtained written consent from them on behalf of their school to take part in the study. We also obtained written consent from the primary-5 (year 5 of primary school) teachers identified by the head teachers. Within each participating school, we included all year-5 children. The official starting age for year-5 children in Uganda is 10 years, but many children are older than this. We did not obtain assent from individual children or consent from their parents. The intervention posed minimal risk and no more risk than other teaching materials, almost none of which have been evaluated. Informed consent by individual children or their parents, in effect, would be meaningless once the decision to participate was taken by the head teacher and the teachers, who have the responsibility and authority to make decisions about lesson plans and the administration of tests. Individual children and their parents had the same right to refuse participation as they do for any other lesson or test in primary schools.

Randomisation and masking

We randomly allocated schools (1:1) to the intervention or control group using a computer-generated sequence with block sizes of four and six and equal allocation ratios within each block. We used stratified randomisation to help ensure equal distribution of schools for two variables: school ownership (public and private) and geographical location (urban, semi-urban, and rural). A statistician who was not a member of the research team, together with his assistants, generated six randomisation lists (one for each combination of the two variables) with unique codes. They labelled opaque envelopes with the unique codes, inserted cards with the study group allocated to each code in the envelopes, and sealed them.

After obtaining consent from 120 schools, two research assistants selected each school from a list of the schools and identified the appropriate randomisation list to be used for that school, based on its geographical location and ownership. They assigned the next available code from that list to each school and then opened the corresponding envelope to determine whether the school was assigned to the intervention or control group. No changes to allocation were made during or after this process.

We informed the participating head teachers and year-5 teachers about the purpose of the study in the consent form (available with the protocol), which they signed before being randomly allocated. After randomisation, they knew whether they were in the intervention or control group. The consent form included information about the outcome measure, stating that it "consists of multiple-choice questions that assess an individual's ability to apply concepts that people must be able to understand and apply to assess treatment claims and to make informed health-care choices". We did not show them the test until the end of the school term. Children in both groups of the trial were informed of the purpose of the test used as the primary outcome measure when their teachers asked them to complete it at the end of the term. Because the teachers and children wanted to know their scores, they put their names on the tests and were told that they and their teachers would be told their scores. The statistician who analysed the data did not know which group was the intervention and control group when he did the primary analyses, but this became obvious due to the magnitude of the effect.

Procedures

We first identified the key concepts that people need to understand and apply when assessing claims about treatments.[14] Together with teachers in Uganda, we established which of those concepts were relevant for primary school children (appendix 1). We started with a list of 32 key concepts, divided into six groups:[15] recognising the need for fair comparisons of treatments, judging whether a comparison of treatments is a fair comparison, understanding the role of chance, considering all the relevant fair comparisons, understanding the results of fair comparisons of treatments, and judging whether fair comparisons of treatments are relevant.

We consulted with Ugandan teachers, who found all six groups of concepts to be relevant for year-5 children. Based on these consultations with the teachers, we judged that 24 of the 32 concepts could be learned by primary school children. These final judgments were made by members of the research team in a face-to-face meeting using informal discussion to reach a consensus.

We developed the resources iteratively between 2013 and 2015, using idea generation and prototyping, pilot testing with non-participatory observation, user-test interviews with children and teachers, and feedback from a network of teachers (appendix 1). We found that there were too many concepts to teach in a single school term. We therefore considered the importance and difficulty of each concept, informed by data from the piloting and user testing. Based on these considerations, we selected 12 concepts (panel 1).

Panel 1: **12 key concepts covered by the Informed Health Choices primary school resources**

Claims
- Treatments might be harmful
- Personal experiences or anecdotes (stories) are an unreliable basis for assessing the effects of most treatments
- Widely used treatments or treatments that have been used for a long time are not necessarily beneficial or safe
- New, brand-named, or more expensive treatments may not be better than available alternatives
- Opinions of experts or authorities do not alone provide a reliable basis for deciding on the benefits and harms of treatments
- Conflicting interests may result in misleading claims about the effects of treatments

Comparisons
- Evaluating the effects of treatments requires appropriate comparisons
- Apart from the treatments being compared, the comparison groups need to be similar (ie, "like needs to be compared with like")
- If possible, people should not know which of the treatments being compared they are receiving
- Small studies in which few outcome events occur are usually not informative and the results may be misleading
- The results of single comparisons of treatments can be misleading

Choices
- Treatments usually have beneficial and harmful effects

The concepts are shown here as they are described in the key concepts list, which was not designed as a learning resource, and not as they were presented to the children in the primary school resources (appendix 1).

The resulting learning resources included a textbook, a teachers' guide, exercise books, a poster, activity cards, and a song. The textbook (appendix 1) of a story told in a comic book format (figure 1), instructions for classroom activities, exercises, a checklist summarising the concepts in the book, and a glossary of key words with definitions in English and translations to Luganda and Swahili. In addition to the textbooks, we provided intervention schools with a guide for each teacher, exercise books for each child, a poster of the checklist for the classroom, and activity cards for the seventh lesson (appendix 1). We also provided them with the "Think carefully about treatments" song on an MP3 player (appendix 1). The lyrics of this song are another reminder of the key messages in the book. Panel 2 lists the contents of the book and the teachers' guide. Appendix 1 provides a description of the intervention using the GREET TIDieR checklist.

There are three school terms per year in Ugandan primary schools, each lasting between 12 and 14 weeks. Teaching periods last 40 min. We designed the resources to be used over 9 weeks, with one double period (80 min) per week, during a single term, and 1 h to complete the test at the end of the term. There was an expectation on the part of the head teachers and teachers that any content displaced by the lessons would be compensated, so that time was not taken away from other lessons. Each school decided how to do this.

At least 1 week before the trial began, and before the introductory workshop, we gave teachers' guides to the teachers in the intervention schools, enabling them to familiarise themselves with the content and prepare a plan for delivering the lessons. We invited all participating teachers in the intervention group to attend a 2 day introductory workshop. At the workshop, we (AN and DS) informed them about the study objectives and procedures, including the general nature of the outcome measure; went through all nine lessons outlined in the primary school resources; and addressed any questions or concerns that arose.

We monitored delivery of the intervention, in accordance with guidelines of the Ministry of Education school supervisory timetable. These allow for follow-up of newly introduced programmes within schools. One of the investigators (AN or DS) or a research assistant observed one lesson in each of the classes in the intervention schools. If there were not enough textbooks, we provided these; if schools were behind schedule in completing the lessons, we explored why; and we addressed any administrative issues relating to the conduct of the trial. We observed how the teachers taught the lessons, but we did not provide feedback or advice to the teachers.

We also encouraged the teachers to make summaries for themselves after reading each chapter in the teachers' guide in preparation for the lesson, and we asked them to hand these in to the study team after the intervention period. We did this to help ensure that the teachers read the teachers' guide in preparation for the lessons, as well as to collect data for the process evaluation.

We contacted the schools allocated to the control group at the beginning of the school term, and invited year-5

Figure 1: **An excerpt from the comic book story in the textbook**

teachers to a 2 h introductory meeting in each district. At these meetings, we informed them about the study procedures, including the general nature of the test that we would be using as the outcome measure. We told them that they would receive the primary school resources at the end of the study. We did not introduce them to the resources or invite them to an introductory workshop.

Children in both groups of the trial completed the test in their classrooms at the end of the term. Research assistants delivered the tests a few hours before exam time and collected them immediately after the exam. They ensured that the children had sufficient time to complete the test (1 h, as is current practice for primary school exams in Uganda). All reading materials, including the Informed Health Choices poster, were removed from the class during exam time. The children (where possible) had spacing that is at least double the usual sitting class spacing, and the test was completed individually without assistance, under supervision of the teachers and observed by the research assistants. Most teachers completed the test at the same time as the children. We contacted teachers who were not available on the day of the exam to arrange completion of the questionnaire on another day. The children and the teachers were aware that missing answers would be scored as wrong.

Outcomes

The primary outcome was measured at the individual participant level as: the mean test score (percentage of correct answers) on the test taken at the end of the term and the proportion of children with a passing score. The secondary outcomes were the proportion of children with a passing score for a subgroup of children who received an audio version of the test in Luganda; the proportion of children with a score indicating mastery of the concepts; for each concept, the proportion of children who answered both questions correctly; the children's intended behaviours and self-efficacy; and the children's attitudes towards science and school. Additionally, we have reported the following, which were not specified in the protocol: mean scores, passing scores, and mastery scores for the teachers, the standardised mean difference for the children, and the cost of the intervention

The test at the end of the term included 24 multiple-choice questions (two for each concept) from the Claim Evaluation Tools database (appendix 1).[16] The questions had between two and four response options, with an overall probability of answering 39% of the questions correctly by chance alone. We developed the questions based on extensive feedback from methodological experts, health professionals, teachers, children, and members of the public.[16] We conducted two Rasch analyses to validate the test.[17,18] Most year-5 school children in Uganda do not have English as their first language and many have poor reading skills. Because we were concerned that this might affect their scores on the test, we also developed a Luganda version of the test to be administered orally to a subgroup of children in each school to estimate the effect of literacy on test scores.[18] We asked the teachers at each school to select 15 children who had already taken the written test in English and who were competent in Luganda. In schools with small classes, the Luganda version was received by all the children who met those two criteria and were present on the day of the oral test.

Two additional multiple-choice questions were included, making 26 in total. These were included because the test used in this trial was also used in a linked randomised trial evaluating a podcast given to the parents of some of the children at the end of the term.[19] These two extra questions addressed the concept: “a treatment outcome may be associated with a treatment, but not caused by the treatment”. This concept was not covered in the primary school resources and responses to the two extra questions were not included in the primary analyses.

The test included questions that assessed intended behaviours, self-efficacy (“an individual’s conviction of their own capability to complete a task or perform a particular behaviour in order to realise goals”), and attitudes (appendix 1). There were four questions that assessed reading skills. We used the answers to those four questions as a covariate in exploratory analyses. In the intervention group, the test included questions that assessed satisfaction with the resources.

We used an absolute (criterion referenced) standard to set a passing score (appendix 1). Children were counted as “passing” or “failing” depending on whether they met this prespecified criterion. We used a combination of Nedelsky’s and Angoff’s methods to determine the cutoff for a passing score. Additionally, using the same methods, we determined a second cutoff for a score that indicated mastery of the 12 concepts. The criterion for passing was a minimum of 13 of 24 questions answered correctly. The criterion for mastery was a minimum of 20 of 24 questions answered correctly.

We will report comparisons of academic achievement using end of term examinations as well as attendance between children in the two groups in the process evaluation in a separate report.[13]

We have reported three additional outcomes that were not specified in the trial protocol: the teachers’ scores on the test, which was planned as part of the process evaluation; the standardised mean difference for the children’s test scores, which allows comparison with effect sizes from other studies; and the cost of the intervention. We estimated the cost of the intervention, based on the actual printing costs, and estimated costs for delivery of the materials, teacher workshops, and teachers’ time. We assumed the teaching materials, apart from the exercise book and the test, would be used over 5 years; the training workshops for the teachers would not need to be repeated during this time; and an interest rate of 5%, giving an annualisation factor of 0·23.

All the outcomes were measured at the end of the school term in which the intervention was implemented. We will measure the sustainability of the effects after 1 year. We asked teachers to record unexpected adverse events and problems that might pose risks to the children or others, and asked them to report these to the investigators or to the Makerere University College of Health Sciences, Institutional Review Board.

Teachers in the intervention group of the trial were given the contact information of the principle investigators (AN and DS) at the start of the trial and instructions for

Panel 2: **Contents of the textbook and the teachers’ guide**

The Health Choices Book: learning to think carefully about treatments, a health science book for primary school children

Introduction
- Lesson 1: Health, treatments, and effects of treatments

John and Julie learn about CLAIMS about treatments
- Lesson 2: Someone’s experience using a treatment
- Lesson 3: Other bad bases for claims about treatments (part 1)
- Lesson 4: Other bad bases for claims about treatments (part 2)

John and Julie learn about COMPARISONS of treatments
- Lesson 5: Comparisons of treatments
- Lesson 6: Fair comparisons of treatments
- Lesson 7: Big enough fair comparisons of treatments

John and Julie learn about CHOICES about treatments
- Lesson 8: Advantages and disadvantages of a treatment

Review
- Lesson 9: Review of what is most important to remember from this book

Teachers’ guide

The teacher’s guide includes an introduction to the project and the resources, and the following for each lesson, in addition to the embedded chapter from the textbook:
- The objective of the lesson
- A lesson preparation plan
- A lesson plan
- A list of materials that the teacher and children will need
- A synopsis of the story
- Keywords in the chapter
- Review questions to ask the children after reading the story
- Extra examples for illustrating the concepts
- Background about examples used in the story
- Teacher instructions for the classroom activity
- Answers and explanations for the activity
- Answers and explanations for the exercises
- Background information, examples, and keyword definitions for teachers

recording adverse events and problems in journals that they were asked to keep for the process evaluation. For the process evaluation, which will be reported separately, we have collected in-depth qualitative data from interviews and focus group discussions regarding participants' views of the intervention, potential adverse effects, as well as other potential benefits of the intervention.

Statistical analysis

We used the University of Aberdeen Health Services Research Unit's Cluster Sample Size Calculator to calculate the sample size, applying the following assumptions: 70 children per cluster; an intraclass correlation coefficient of 0·5, based on ICCs from a meta-analysis of randomised trials of school interventions and an international comparison of ICCs for educational achievement outcomes, which suggested the ICC might be very high;[20,21] 0% as the proportion of children expected to achieve a passing score without the intervention, based on findings from pilot testing; 10% as the smallest difference we wanted to be able to detect; an alpha of 0·05; and a power of 90%.Based on these assumptions, we estimated that we would need 50 schools in each group. Allowing for a loss to follow-up of up to 10% (for schools where it might be impossible to administer the tests at the end of the term), we estimated that we needed a minimum of 55 schools in each group.

For the primary and secondary outcomes, we used mixed models with a random effects term for the clusters and the stratification variables modelled as fixed effects, using generalised logistic regression for dichotomous outcomes and linear regression for continuous outcomes. The statistical analyses were done with R (R Core Team, Vienna, Austria; version 3.3.2). All the children and teachers who completed the test were included in the analyses. Missing values were counted as wrong answers. We converted odds ratios from logistic regression analyses to adjusted differences using the intervention group percentage as the reference for the main outcomes and the control group percentage as the reference for the secondary outcomes.

See Online for appendix 2

We did two post-hoc sensitivity analyses suggested by external reviewers to explore the risk of bias due to attrition, which was larger in the control schools than in the intervention schools. First, we did a weighted analysis using inverse probability weighting. In this analysis, the children in each school were given a weight equal to the inverse of the proportion of children in the school that completed the test. Second, we calculated upper and lower bounds for the mean difference in test scores using the Lee bounds approach.[22] These are constructed by trimming the group with less attrition at the upper and lower tails of the outcome (test score) distribution respectively. In this analysis, the sample was trimmed in the intervention schools so that the proportion of children included in the analysis was equal for both groups. We did not adjust for covariates in this analysis.

For each outcome, we have reported the proportion, mean and standard deviation or count and percentage for each group, the estimated difference, the estimated confidence interval for the difference, and the p value from the statistical models. For questions about intended behaviours and self-efficacy, we dichotomised the responses in the analysis (eg, very unlikely or unlikely *vs* very likely or likely), and reported the number and percentage of children for each of the response options.

Based on data from the pilot studies, we anticipated that many of the children would have poor reading skills, and that this might impede their ability to comprehend the content of the textbook and to answer the multiple-choice questions. We explored whether there were differences in the effect of the intervention for children with advanced reading skills (all four literacy questions answered correctly) versus basic reading skills (both basic literacy questions correct and one or two of the advanced literacy questions wrong) versus lacking basic reading skills (one or both basic literacy questions wrong).

We calculated the adjusted standardised mean difference (Hedges' g) so that we could put the effect of the intervention in the context of effect size reported for other interventions to improve critical thinking or learning in primary schools.[11,20,23] We calculated an adjusted Hedges' g and its 95% confidence interval using formulae described by White and Thomas.[24]

We intended to do a second subgroup analysis to explore whether having a parent who listened to the podcast improved the scores of the children and whether there was an interaction between the effect of the podcast and the primary school resources. However, because of delays in starting the podcast trial, the parents allocated to listen to the podcast did not do so until after the children had completed the tests. There was no data monitoring committee. Appendix 2 provides data files for the study.

This trial is registered with the Pan African Clinical Trial Registry, number PACTR201606001679337.

Role of the funding source

The funder had no role in study design, data collection, data analysis, data interpretation, or writing of the report. The principal investigator (AN) had full access to all the data in the study and had final responsibility for the decision to submit for publication.

Results

Between April 11, 2016, and June 8, 2016, 2960 schools in Uganda were assessed for eligibility. After recruitment meetings, 120 schools consented and were randomly assigned to either the intervention (n=60) or control group (n=60). All 120 schools provided data and were included in the analysis. Figure 2 shows the

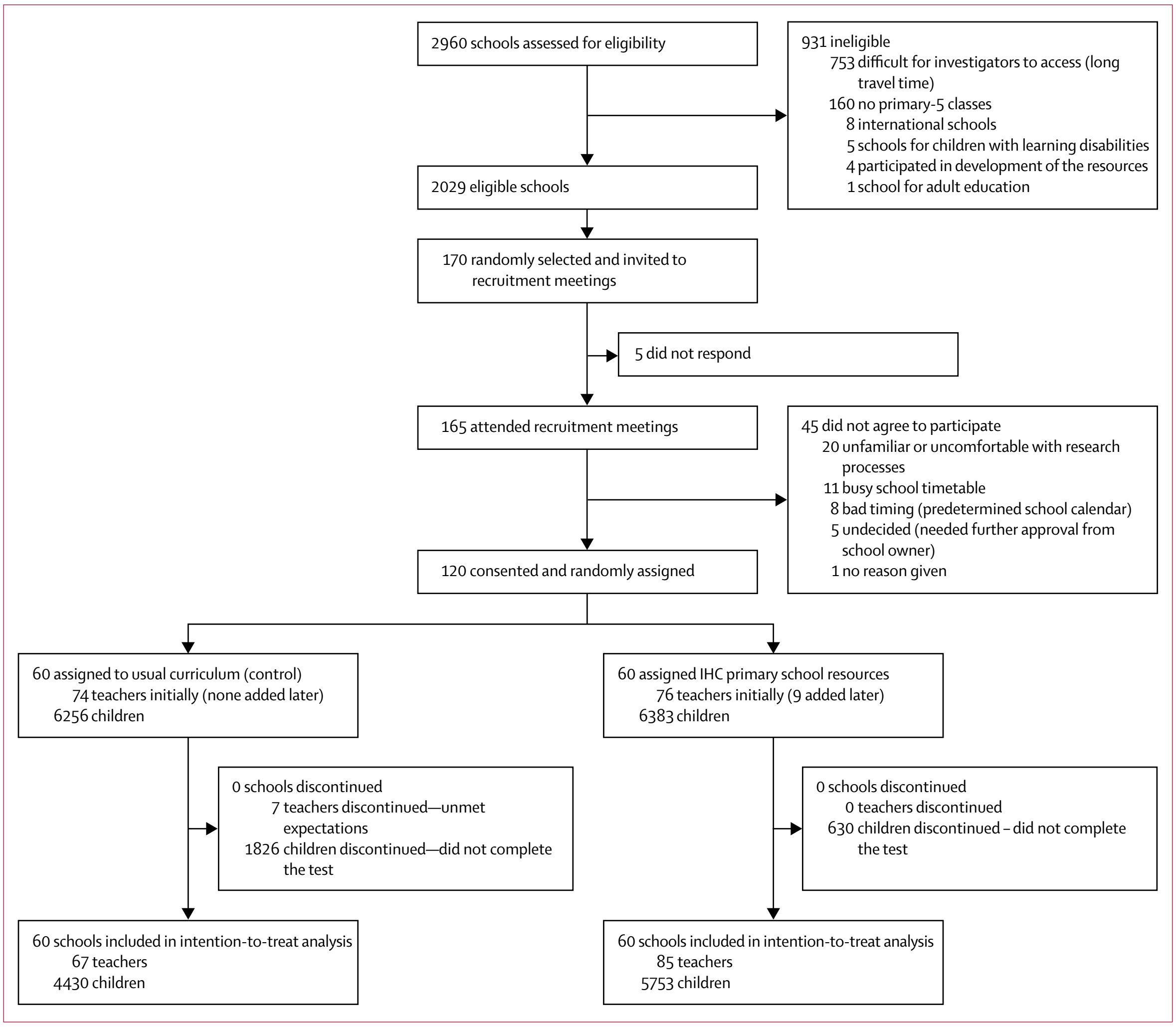


Figure 2: **Trial profile**

reasons for non-inclusion, the flow of the schools, teachers, and children through the study. Most of the schools in both the intervention and control groups were urban or semi-urban (table 1). There were more public schools in the control group (55% *vs* 50%). There were more teachers with a university degree and fewer with a teaching diploma in the intervention schools (12% *vs* 5% and 41% *vs* 49%), and there were more teachers who taught science as their main subject in the intervention group (80% *vs* 73%; table 1). These minor differences seem unlikely to have biased the results. In the sensitivity analysis, the Luganda test was administered in 114 schools. Six control schools declined (five because of a lack of time, and one because no children reportedly spoke Luganda).

All 60 schools in the intervention group delivered all nine of the lessons. The timing of the lessons varied. Some schools (mostly boarding schools) did the lessons early in the morning or in the evening. Others taught the lessons when the usual science lessons were scheduled or when co-curricular activities like drama and sports were scheduled. These schools compensated for what was missed by doing the missed activities early in the morning or in the evening.

We initially asked each head teacher to select one year-5 science teacher, but some schools had more than one teacher who taught year-5 science, so there were more than 60 teachers in both the intervention and control schools. Six intervention schools that had more than one year-5 class (with a different teacher for each class) identified altogether nine more teachers for whom they requested training. No teachers were added in the control schools, since the teachers in the control schools did not receive training. All 85 teachers in the intervention schools and 67 (91%) of the teachers in the control schools completed the same test that the children took at the end of the term.

Altogether, 10 183 children completed the test. More children completed the test in the intervention schools (5753 [90%] of 6383) than in the control schools (4430 [71%]

of 6256). This was most likely because teachers in the intervention schools were more motivated to request that the children stay at the end of the term to take the test, having committed time and energy to the intervention, than teachers in the control schools, who taught the usual curriculum. There was no appreciable difference in the proportion of girls (45%) or the median age of children in the two comparison groups (11 years, IQR 10–12). Most of the children answered all the questions. The proportion of missing values (unanswered questions) for each question was between 0·5% and 4·3% and the number of missing values was similar between the intervention and control schools (p=0·964; appendix 1).

The average score for children in the intervention schools was 62·4% (SD 18·8) compared with 43·1% (15·2) in the control schools. The adjusted mean difference (based on the regression analysis) was 20·0% (95% CI 17·3–22·7; p<0·00001) higher in the intervention than in the control group. Appendix 1 shows the distribution of test scores. In the intervention schools, 3967 (69%) of 5753 children had a passing score (≥13 of 24 correct answers), compared with 1186 (27%) of 4430 in the control schools (table 2). The adjusted difference (based on the odds ratio from the logistic regression analysis) was 50% more children who passed (95% CI 44–55; p<0·00001) in the intervention than in the control group.

The average score for the 1616 children who completed the test orally in Luganda was 66·3% in the intervention schools compared with 49·7% in the control schools. The adjusted difference was 15·8% (95% CI 12·7–19·0), which was slightly smaller than the adjusted mean difference for the written test (table 3). We did two additional sensitivity analyses to assess the potential risk of bias from attrition—ie, children who did not take the test. There was very little difference between the results of the weighted analysis, using inverse probability weighting, and the primary analysis (table 3), suggesting that the results are robust. In the second analysis, we calculated Lee bounds for the mean difference in test scores. This resulted in a lower (worst case) and upper (best case) mean difference of 14·2% and 24·6%, respectively (95% CI 13·4–25·5). This indicates that even with the worst-case scenario, the average test score in the intervention schools was still 14·2% higher than in the control schools (with a lower confidence limit of 13·4%). Moreover, the worst-case scenario, which removed 19% of the children with the highest test scores from the intervention group, is unlikely. This is equivalent to assuming that the children in the control schools who did not take the test would have had scores that corresponded to the top 19% of the children in the intervention schools, had they taken the test. Attrition for each strata of school (based on ownership and location) and test scores for each stratum are summarised in appendix 1.

In the intervention schools, 19% of the children had a score indicating mastery of the 12 key concepts (≥20 of 24 correct answers) compared with 1% of the children in the control schools. The adjusted difference was 18% more children in the intervention schools who mastered the concepts (95% CI 18–18; p<0·00001).

	Control schools	Intervention schools
Schools		
Schools (selected from the central region of Uganda)	60	60
Location		
Rural	8 (13%)	6 (10%)
Semi-urban	15 (25%)	14 (23%)
Urban	37 (62%)	40 (67%)
Ownership		
Public	33 (55%)	30 (50%)
Private	27 (45%)	30 (50%)
Teachers		
Teachers (initially identified by head teachers)	74	76
Completed tests*	67 (91%)	85 (100%)†
Education‡		
Certificate	30 (45%)	39 (46%)
Diploma	33 (49%)	35 (41%)
University degree	3 (4%)	10 (12%)
Main subject taught		
Science	49 (73%)	68 (80%)
Sex		
Women	29 (43%)	34 (40%)
Children		
Children (enrolled in year-5 at the start of the term)	6256	6383
Completed tests*	4430 (71%)	5753 (90%)
Median completed tests per class§	60 (40–95)	61 (43–89)
Sex		
Girls	1973 (45%)	2599 (45%)
Median age (years)	11 (10–12)	11 (10–12)

Data are n, n (%), or median (IQR). *Questions about the characteristics of the teachers and children were included in the test completed at the end of the school term. †Head teachers were initially asked to identify teachers who taught science to children in the fifth year of primary school. However, some schools had more than one year-5 class. Six intervention schools with more than one year-5 class (with a different teacher for each class) requested that nine additional teachers be included altogether. ‡There was one missing value in each group for this variable. §The average class size at the start of the term was 84 children in both groups.

Table 1: **Characteristics of the participants**

For each concept, the proportion of children who answered both questions correctly was higher in the intervention schools than in the control schools, including for the concept that was not covered in the primary school resources (p<0·00001 for all 13 concepts after a Bonferroni correction for multiple comparisons; figure 3).

Children in the intervention schools were more likely to respond that they would find out what a claim was based on (adjusted difference 10·6%, 95% CI 6·2–14·7); find out if a claim was based on research (10·8%, 6·3–15·1); and participate in a research study if asked

	Control schools (60 schools, 4430 children)	Intervention schools (60 schools, 5753 children)	Adjusted difference*	Odds ratio†	ICC
Primary outcome					
Mean score	43·1% (15·2)	62·4% (18·8)	20·0% (17·3–22·7)	..	0·18
Children with a passing score (≥13 of 24 correct answers)	1186 (27%)	3967 (69%)	50% (44–55)	9·3 (6·6–13·2)	0·19
Secondary outcome					
Children with a mastery score (≥20 of 24 correct answers)	38 (1%)	1070 (19%)	18% (18–18)	35·3 (20·6–60·7)	0·21
Teachers' scores‡					
Mean score	66·7% (14·3)	84·6% (17·1)	18·3% (12·9–23·3)	..	..
Teachers with a passing score (≥13 of 24 correct answers)	58 (87%)	83 (98%)	11% (4–13)	7·2 (1·5–35·3)	..
Teachers with a mastery score (≥20 of 24 correct answers)	10 (15%)	61 (72%)	57% (37–70)	14·4 (6·2–33·1)	..

Data are % (SD), % (95% CI), or n (%). ICC=intraclass correlation coefficient. *The adjusted difference is based on mixed models with a random-effects term for the clusters and the stratification variables are modelled as fixed effects, using logistic regression for dichotomous outcomes and linear regression for continuous outcomes. p<0·0001 for all four comparisons. † The odds ratios from the logistic regressions for passing scores and mastery scores have been converted to differences based on the intervention school proportions and the odds ratios calculated using the intervention schools as the reference (the inverse of the odds ratios shown here). ‡There were 67 teachers in the control schools and 85 in the intervention schools.

Table 2: **Main results**

(7·8%, 3·7–11·9), compared with children in the control schools (appendix 1).

Children in the intervention schools were more likely to consider it easy to assess whether a claim is based on research (adjusted difference 15·0%, 95% CI 10·9–19·0) compared with children in the control schools (appendix 1). They were less likely to consider it easy to assess how sure they could be about research results (adjusted difference –4·1%, 95% CI –1·0 to –7·3). We detected little if any difference in how easy they thought it was to find information about treatments based on research, or to assess how relevant research findings are likely to be to them. We also detected little if any difference in attitudes towards school or science. At least 90% of the children overall indicated a positive attitude in response to all four questions (appendix 1). Most children (4864 [85%] of 5753) in the intervention schools had positive views of the textbook (appendix 1).

None of the teachers or research assistants who observed the lessons reported any adverse events. Although the intervention had positive effects regardless of reading skills (appendix 1), there was an interaction between levels of reading skills and the effects of the intervention. As we hypothesised, the beneficial effects of the intervention were larger for children with better reading skills.

In an analysis that was planned for the process evaluation, but not included in the protocol for the trial, we found that most teachers in both the control and the intervention groups (87% and 98%, respectively) had a passing score on the same test that the children took at the end of the term (adjusted difference 11%, 95% CI 4–13; table 2). The teachers in the intervention group were much more likely to have a score indicating mastery of the concepts (72% *vs* 15%; adjusted difference 57%, 95% CI 37–70).

	Adjusted difference*	SMD† or OR
Primary analysis	20·0% (17·3–22·7)	SMD 1·16 (1·00–1·32)
Weighted analysis	20·0% (17·3–22·7)	SMD 1·08‡ (0·93–1·22)
Lee bounds	14·2–24·6% (13·5–25·5)	..
Oral examination in Luganda§	15·8% (12·7–19·0)	SMD 0·99 (0·79–1·20)
Passing score (≥13 out of 24 correct answers)		
Primary analysis	49·8% (43·8–54·6)	OR 9·3 (6·6–13·2)
Weighted analysis	50·0% (44·1–54·8)	OR 9·5 (6·7–13·4)

Data are % (95% CI) unless stated otherwise. SMD=standardised mean difference. OR=odds ratio. *The adjusted difference is based on mixed models with a random-effects term for the clusters and the stratification variables modelled as fixed effects, using logistic regression for dichotomous outcomes and linear regression for continuous outcomes. p<0·0001 for all analyses. The ORs from the logistic regressions for passing scores have been converted to differences based on the intervention school proportions and the ORs calculated using the intervention schools as the reference (the inverse of the ORs shown here). †Adjusted Hedges' g. ‡The effect size is different from the primary analysis, despite the adjusted mean difference being the same, because of a difference in the intraclass correlation coefficient. §Administered to 769 children in the control schools (mean 49·7% [SD 15·6]) and 847 children in the intervention schools (66·3% [15·7]).

Table 3: **Sensitivity analyses**

We calculated the effect size (standardised mean difference) for the children for comparison with other studies (table 3). The effect size (Hedges' g) was 1·16 (95% CI 1·00–1·32) based on the primary analysis. It was slightly less (1·08; 95% CI 0·93–1·22) based on the weighted analysis. We estimated that the average annual

	Control schools % correct* (n=273)	Intervention schools % correct* (n=288)	Adjusted difference%† (95% CI)	Difference* per 1000	Odds ratio (95% CI)	ICC‡
Claims						
Treatments might be harmful	21·0 (n=930)	52·1 (n=2999)	36·1 (28·6–42·9)	361 more	5·0 (3·7–6·7)	0·12
Personal experiences or anecdotes (stories) are an unreliable basis for assessment of the effects of most treatments	14·5 (n=643)	41·5 (n=2387)	31·4 (25·0–37·0)	314 more	5·0 (3·8–6·3)	0·11
A treatment outcome might be associated with a treatment, but not caused by the treatment§	6·9 (n=306)	28·3 (n=1630)	21·2 (16·7–27·7)	212 more	5·3 (4·2–7·1)	0·09
Widely used treatments or treatments that have been used for a long time are not necessarily beneficial or safe	16·0 (n=708)	40·6 (n=2335)	30·4 (23·6–38·3)	304 more	4·5 (3·4–6·3)	0·15
New, brand-named, or more expensive treatments might not be better than available alternatives	33·1 (n=1467)	62·7 (n=3609)	33·3 (27·6–39·2)	333 more	4·0 (3·1–5·3)	0·11
Opinions of experts or authorities do not alone provide a reliable basis for deciding on the benefits and harms of treatments	27·1 (n=1201)	51·2 (n=2944)	27·4 (21·7–33·7)	274 more	3·2 (2·6–4·2)	0·10
Conflicting interests might result in misleading claims about the effects of treatments	19·8 (n=875)	34·6 (n=1990)	15·1 (10·4–19·6)	151 more	2·2 (1·8–2·6)	0·07
Comparisons						
Evaluation of the effects of treatments requires appropriate comparisons	6·9 (n=306)	28·3 (n=1630)	21·2 (16·7–27·7)	212 more	5·3 (4·2–7·1)	0·10
Apart from the treatments being compared, the comparison groups need to be similar (ie, like needs to be compared with like)	14·6 (n=640)	34·5 (n=1987)	20·1 (14·7–25·8)	201 more	3·1 (2·4–4·0)	0·09
If possible, people should not know which of the treatments they are receiving‡	19·8 (n=877)	34·6 (n=1988)	17·2 (12·4–22·3)	172 more	2·4 (1·9–2·9)	0·06
Small studies in which few outcome events occur are usually not informative and the results may be misleading	21·9 (n=970)	42·5 (n=2445)	22·6 (17·6–27·3)	226 more	2·9 (2·3–3·4)	0·07
The results of single comparisons of treatments can be misleading	22·6 (n=1000)	37·5 (n=2155)	19·0 (12·9–25·1)	190 more	2·4 (1·9–3·1)	0·10
Choices						
Treatments usually have beneficial and harmful effects	27·1 (n=1201)	41·1 (n=2364)	16·0 (10·4–22·4)	160 more	2·0 (1·6–2·6)	0·10

95% CI
95% CI for key concept not included in learning resources
Point estimate

Figure 3: **Results for each key concept**
*There were two multiple-choice questions for each concept. The proportions are for the percentage of children who answered both questions correctly. †The adjusted difference is based on mixed models with a random effects term for the clusters and the stratification variables modelled as fixed effects, using logistic regression. All the p values are less than 0·0001 after being adjusted for multiple comparisons. The odds ratios from the logistic regressions have been converted to differences based on the control school proportions and the odds ratios shown here. ‡Intraclass correlation coefficient. §This concept was not included in the learning resources or counted in the average, pass, or mastery scores.

cost of the intervention, including teachers' time, would be approximately US$400 per school, and $4 per child (appendix 1).

Discussion

Use of the Informed Health Choices primary school resources had a large effect on the ability of primary school children in Uganda to assess claims about treatment effects. This effect was larger for children with better reading skills, but the intervention was effective for children lacking basic reading skills, as well as for children with basic or advanced reading skills. This effect was achieved even though the learning materials and the tests were in English, which was not the children's first language. Based on findings from pilot testing both the resources and the test used to measure the outcomes, we were surprised by the size of the effect, which is also large in comparison to other education interventions in primary schools in low-income and middle-income countries,[20] and other interventions to teach critical thinking for all ages in high-income countries.[11] In addition, the intervention had a positive effect on the children's intended behaviours and the teachers' mastery of the key concepts.

We have not found any directly comparable studies. Other interventions in primary schools have been found to have a positive effect on critical thinking, but these studies have been conducted in high-income countries and neither the interventions nor the outcome measures are directly comparable.[11] Nonetheless, the effect size for this study (a standardised mean difference of 1·16) is well above the average effect size reported for other critical thinking interventions (0·33 [SD 0·55]; appendix 1).[11] It is larger than any of the effect sizes reported in a systematic review of interventions to improve learning in primary schools in low-income and middle-income countries for interventions with teacher training and for interventions with instructional materials.[20] It is also larger than the effects reported in a second systematic review for learning outcomes for structured pedagogy programmes in low-income and middle-income countries.[23] However, most of those studies used reading or maths tests as the outcome measure rather than a test that was explicitly designed to measure skills that were the focus of the intervention. Only two of the studies of structured pedagogy programmes measured cognitive or problem solving skills. Therefore, it is not appropriate to compare our results with the studies in these reviews.

The systematic review of interventions to improve learning in primary schools in low-income and middle-income countries found that instructional materials alone may not improve learning, and that they are more likely to be effective when combined with teacher training and a well articulated instructional model.[20] However, the second systematic review of structured pedagogy programmes, all of which included teacher training and many of which provided learning resources, found a large range of effects.[25] Possible explanations for a lack of effect in some studies, identified by the investigators, include teachers sometimes not being knowledgeable or experienced enough to fully understand their training or not implementing the lessons as intended or as often as planned. Another possible explanation was that the investigators did not consider key contextual factors, such as limited resources and high student-to-teacher ratios.[23]

It is uncertain how effective the Informed Health Choices primary school resources would be without the teacher training and support from the school authorities and teachers. The more than 85% of teachers in the control schools (without training) who had passing scores on the test used as an outcome measure suggests that the teachers were knowledgeable enough to understand the training. That 72% of the teachers in the intervention schools had scores indicating mastery of the concepts, compared with 15% of the teachers in the control schools, suggests that the training, together with their teaching experience during the term, was effective. Over 2 years of pilot and user-testing the learning resources, and collaborating with a network of teachers, helped to ensure that our intervention took account of contextual factors, including large student-to-teacher ratios, crowded classrooms, and scarce resources.

No adverse events were reported by any of the head teachers, teachers, children, or parents. Potential adverse effects that were hypothesised before the trial, but were not observed, are summarised in appendix 1. These will also be explored further in the process evaluation.

A limitation of this study is the number of children that did not take the test used to measure outcomes at the end of the term and the difference in the proportion of children that completed the test in intervention schools (90%) and control schools (71%). Attrition is a common problem in randomised trials of education interventions.[20,25] The most likely reason for the difference in attrition in this study is that, having invested time and energy in the lessons, teachers in the intervention schools put more effort into making sure that children in their classes completed the test.

Our study does not meet the attrition standard suggested by the What Works Clearinghouse (WWC).[26] However, that standard is based on tolerating a maximum bias of 0·05 standard deviations, and it is highly sensitive to the maximum level of bias that a systematic review is willing to accept.[27] The effect size for this study (1·16) is more than 20 times the WWC maximum tolerable bias. Although we cannot rule out some degree of bias due to attrition, it is highly unlikely that bias modified the observed effect substantially relative to the size of the effect. The sensitivity analyses that we did support this conclusion (table 3).

There were also more teachers who completed the test in the intervention schools. This was probably because although we initially asked the head teachers to identify one year-5 teacher, some schools had more than one class. We subsequently included all the teachers who taught science to a year-5 class in the intervention schools, but not in the control schools.

Another limitation of this study is that the test used as the outcome measure was aligned with the intervention ("treatment-inherent"). That is, the test measured the ability to apply the concepts that the resources were designed to teach. Treatment-inherent outcome measures are associated with larger effect sizes than independent measures.[28] It is also problematic to compare the effect size from this study with studies in which both comparison groups were taught the subject being tested. Because of this, it is inappropriate to compare the effect of our intervention on our outcome measure to the effects of other interventions on independent measures, such as reading or maths tests. Similarly, one should be cautious when comparing our results to the effects of other interventions to teach critical thinking. The systematic review of critical thinking interventions, noted above, found larger effects for outcome measures developed by one or more of a study's authors for use in the study (0·65, 95% CI 0·52–0·78) than for well established measures of critical thinking (0·40, 0·26–0·53).[11]

Because there was no pre-existing outcome measure suitable for our study,[14] we used an outcome measure that was developed by us for this study.[16–18] However, we used multiple-choice questions from a database of questions that independent research methodologists judged to have face validity, and end-users judged to be relevant and acceptable;[18] we validated the test in two Rasch analyses;[17,18] and a group of independent judges determined the cutoff scores for passing and mastery scores. The multiple-choice questions were designed to require critical thinking on the part of the test-takers and could not be answered by simply repeating content from the learning resources (appendix 1). We were careful to ensure that the examples used in the questions were different from those used in the learning resources, and that the children would be able to understand the language that was used without having used the resources. Neither the teachers nor the children were shown the test or similar multiple-choice questions before taking the test.

What the long-term effects of using the Informed Health Choices primary school resources are; whether they will have an effect on actual health choices and outcomes; whether they will have an effect on other measures of academic achievement; and how transferable the findings of this study are to other countries remain uncertain. We will measure the effects on standardised end-of-term examinations in a process evaluation. We will also measure outcomes again after 1 year. This will provide some indication of the degree to which the learning is sustained. Although we measured intended behaviours, it was not possible to measure actual health choices. We will explore the effects on actual choices when we measure outcomes after 1 year, but this will still be limited since most of the children will not be making many of their own health choices, and their choices will be self-reported.

We have piloted and user-tested an earlier version of the resources in Kenya and Rwanda, and we will pilot and user-test translated versions of the current version of the resources in those countries in 2017. User-testing and trials in other countries are needed. The cost of the intervention (approximately \$4 per child) is substantial in light of government expenditure per primary school student (\$29·4) and estimates of the direct costs of primary school education in Uganda.[9,29] We will explore ways of scaling up the use of the intervention in the process evaluation. Together with school authorities, we will try to find ways of covering the costs of scaling up use of the resources in Uganda.

In addition to the inherent educational value of the resources, there are three arguments for considering using these learning resources or similar approaches to teach these skills to primary school children.

First, low health literacy is consistently associated with poor use of health services and poor health outcomes.[2] Improving critical health literacy is likely to improve those outcomes, even though it is uncertain what if any effect use of these resources alone will have on health outcomes. Second, whether the effect on learning is sustained or not, it would be desirable to reinforce what was learned and to introduce additional key concepts, building on what was learned. Use of these resources should be viewed as a first step in a spiral curriculum (appendix 1). It is important to introduce these key concepts at a young age to lay a foundation for future learning and to reduce the development of misconceptions that become resistant to change later.[10] Third, teaching critical thinking is likely to have a positive effect on academic achievement, in addition to its direct effect on critical thinking skills.[11,12] Teaching critical thinking in connection with claims about treatments engages both children and teachers. As noted by a girl in an international school that piloted an earlier version of the learning resources: this is about "things we might actually use instead of things we might use when we are all grown up and by then we'll forget". An illustration of this was provided by a girl in another class at the same school: "When I was grocery shopping with mom, mom was like, 'Buy this toothpaste! It's new and it's really good!' I looked at another one and it was exactly the same, so I actually bought the cheaper one."

In summary, we believe we have shown reliably that it is possible to teach critical appraisal of treatment claims on a large scale in a low-income country. We have not compared our approach to another because, as far as we are aware, there is currently no other evaluated approach for doing this.[14,30] We believe that the Informed Health Choices primary school resources are an important first step towards enabling children to make informed health decisions as they grow older, as patients, future health professionals, citizens, and future policy makers.

Contributors

AN and DS were the principal investigators. They drafted the protocol with help from the other investigators, and were responsible for the day-to-day management of the trial. NKS and ADO had primary responsibility for overseeing the trial. All the investigators reviewed the manuscript, provided input, and agreed on this version. MO and SR had primary responsibility for developing the primary school resources. AM shared primary responsibility for developing the teachers' guide. All the investigators other than YD contributed to the development of the resources and to the protocol. AA-D had primary responsibility for developing and validating the outcome measure. AN and DS had primary responsibility for data collection. YD did the statistical analysis. The Norwegian Institute of Public Health, recipient of the grant from the Research Council of Norway, is the coordinating centre for the Informed Health Choices project. ADO, SR, AA-D, and IC were principal members of the coordinating group for the trial and, together with NKS and the principal investigators, acted as the steering committee for the trial. They were responsible for final decisions about the protocol and reporting of the results.

Declaration of interests

We declare no competing interests.

Acknowledgments

The trial was funded by the Research Council of Norway, Project number 220603/H10. We are grateful for support for this research from the Global Health and Vaccination Research (GLOBVAC) programme of the Research Council of Norway, Developing Excellence in Leadership, Training and Science (DELTAS) African Academy of Science (AAS)/ Alliance for Accelerating Excellence in Science in Africa (AESA) for supporting NKS and DS (DELTAS grant DEL-15-011), and to the English National Institute for Health Research for supporting IC and the James Lind Initiative. Alun Davies, Martha Gerrity, Lena Nordheim, Peter O Okebukola, Newton Opiyo, Jonathan Sharples, Helen Wilson, and Charles Shey Wiysonge determined the cut off scores for passing and mastery. Miriam Grønli was responsible for the textbook colouring and Nora Rosenbaum assisted. We are grateful to Philip Abrami and his colleagues for permission to use a figure from their meta-analysis. Aisha Hashi, Sara Jaber, Rida Shah, and Katie Tveiten helped test prototypes. Michael Mugisha, Anne-Marie Uwitonze, and Jenny Moberg helped with piloting and user testing an earlier version of the learning resources. Carl Lombard, Patrick McEwan, Jonathan Sharples, Helen Wilson, Pam Hanley, Steven Higgins, Alun Davies, Mike English, and Pablo Alonso Coello provided helpful comments on an earlier version of this report. We want to thank Daniel Nkaada at the Ugandan Ministry of Education for technical guidance; Sarah Natunga at the National Curriculum Development Centre in Uganda for reviewing the materials; Margaret Nabatanzi, Solomon Segawa, Martin Mutyaba, and Esther Nakyejwe for their help with data management; and all the research assistants who helped with data collection and entry. We would also like to thank the Informed Health Choices advisory group. We are especially grateful to the many teachers and children in Uganda, Kenya, Rwanda and Norway who helped with the development of the Informed Health Choices primary school resources, and to all the children, teachers, and head teachers at the following schools who participated in this trial: Banda Church of Uganda Primary School, Basiima Primary School, Bat Valley Primary School, Bilal Islamic Primary School, Budo Junior School, Buganda Road Primary School, Bunga Hill Primary School, Busawula Primary School, Buwaate Church of Uganda Primary School, Bweyogerere Church of Uganda Primary School, C & D Primary School, Nansana, Caring Heart Primary School, Childcare Academic Primary School, Kakiri, Cilia Junior School, East Kololo Primary School, Gayaza Junior School, Glory International School, Holy Family Primary School, Homisdallen Primary School, Gayaza, Homisdallen Primary School, Kamwokya, House of Joy Primary School, Kalagi, Jinja Kaloli Primary School, Joy Primary School, Salaama, Kampala Quality Primary School, Kanyange Primary School, Kasana Quality Primary School, Luweero, Katikamu Kisule Primary School, Kawempe Junior School, Kawempe Muslim Primary School, Kawuku Parents' School, Kazo Church of Uganda Primary School, Kazo Mixed Primary School, Kazo Summit Primary school, Kampala Capital City Authority Primary School Kamwokya, Kampala Capital City Authority Primary School, Busega, Kibiri Catholic Primary School, King's Way Primary School, Kireka Seventh Day Adventist Primary School, Kisugu Preparatory School, Kisuule Primary School, Kitemu Integrated Primary School, Kitende Primary School, Kyaliwajjala Uganda Muslim Education Association Primary School, Luweero Seventh Day Adventist Primary School, Maganjo Uganda Muslim Education Association Primary School, Makerere Church of Uganda Primary School, Mbuya Church of Uganda Primary School, Melody Junior School, Kawempe, Mother Kevin Memorial School, Mukono Boarding Primary School, Mukono Town Muslim Primary School, Munyonyo Parents' School, Mutundwe Church of UgandaPrimary School, Mutungo Kitiko Church of Uganda Primary School, Najjanankumbi Young Choir Primary School, Nakasero Primary School, Nakivubo Blue Primary School, Nakyessanja Primary School, Namaliga Church of Uganda Primary School, Namasuba Uganda Muslim Education Association Primary School, Namilyango Junior Boys' School, Namirembe Infants Primary School, Namirembe Parents School, Nampunge Primary School, Namugongo Model Primary School, Namulonge Church of Uganda Primary School, Namulonge Community Primary School, Namungoona Orthodox Primary school, Nansana Church of Uganda Primary School, Nansana Catholic Primary School, Nansana Central Academy Primary School, Nansana Kissa Primary School, Nansana Seventh Day Adventist Primary School, Nanziga Public Primary School, Ndejje Junior School, New Era Primary School, Kawempe, Nkumba Primary School, Nkumba Quran Primary School, Nsangi Mixed Primary School, Outspan Primary School, Namulonge, Pearl Primary School, Kakiri, Pioneer Primary School, Namasuba, Precious Gift Nursery and Primary School, Real Quality Junior School, Sam Iga Memorial Primary School, Seguku Primary School, Sir Apollo Kaggwa Primary school, Nakasero, St Aloysious Primary School, Kyengera, St Anne Girls' Primary School, Naddangira, St Charles Primary School, Bukerere, St Dominic Primary School, Kigo Lunya, St Edward, Primary School, Gobero, St Francis Day and Boarding Primary School, Ntinda, St James Primary School, Nkumba, St Joseph Primary School, Maya, St Jude Primary School, Naguru, St Kizito Primary School, Bugolobi, St Kizito Primary School, Kisozi, St Kizito Primary School, Mpala, St Lawrence Junior School, Kabowa, St Mary's Junior School, Nabbingo, St Pius Mixed Primary School, Naddangira, St Pius Primary School, Masajja, St Theresa Girls' Primary School, Kisubi, St Thomas Bazadde Primary School, Bweyogerere, Step by Step Junior School Kisimu, Summer Junior School, Kawempe, Tendo Primary School, Train Up a Child Primary School, Tripple 'P' Primary School, Ttula Primary School, Victor Junior School Nabbingo, Victorious Infant Primary School, Mukono, Victorious Primary School,Bakuli, Victors' Junior Christian School, Winston Standard Academy Primary School, Wisdom Primary School, Lugala, Wobulenzi Public School, Yudesi Primary School, Zamzam Islamic Primary School.

References

1 Hoffmann TC, Del Mar C. Patients' expectations of the benefits and harms of treatments, screening, and tests: a systematic review. *JAMA Intern Med* 2015; **175:** 274–86.

2 Berkman ND, Sheridan SL, Donahue KE, Halpern DJ, Crotty K. Low health literacy and health outcomes: an updated systematic review. *Ann Intern Med* 2011; **155:** 97–107.

3 Frass M, Strassl RP, Friehs H, et al. Use and acceptance of complementary and alternative medicine among the general population and medical personnel: a systematic review. *Ochsner J* 2012; **12:** 45–56.

4 Brownlee S, Chalkidou K, Doust J, et al. Evidence for overuse of medical services around the world. *Lancet* 2017; published online Jan 8. http://dx. doi.org/10.1016/S0140-6736(16)32585-5.

5 Starr RR. Too little, too late: ineffective regulation of dietary supplements in the United States. *Am J Public Health* 2015; **105:** 478–85.

6 National Research Council. Taking science to school: learning and teaching science in grades K-8. Washington, DC: The National Academies Press, 2007.

7 Sandoval WA, Sodian B, Koerber S, Wong J. Developing children's early competencies to engage with science. *Educ Psychol* 2014; **49:** 139–52.

8 Ruddock G, Sainsbury M. Comparison of the core primary curriculum in England to those of other high performing countries. Research Report DCSF-RW048. London: Department for Children, Schools and Families, 2008.

9 UNESCO Institute for Statistics. UIS.STAT. December 2016 release. http://data.uis.unesco.org/index.aspx?queryid=192 (accessed Jan 16, 2017).

10 Vosniadou S. International handbook of research on conceptual change. 2nd edition. Oxford: Routledge, 2013.

11 Abrami PC, Bernard RM, Borokhovski E, Waddington DI, Wade CA, Persson T. Strategies for teaching students to think critically a meta-analysis. *Rev Educ Res* 2015; **85:** 275–314.

12 Hanley P, Slavin R, Elliott L. Thinking, doing , talking science: evaluation report and executive summary. London: Education Endowment Foundation, 2015. https://v1.educationendowmentfoundation.org.uk/uploads/pdf/Oxford_Science.pdf (accessed March 8, 2017).

13 Nsangi A, Semakula D, Glenton C, et al. Resources to teach primary school children in low-income countries to assess claims about treatment effects: protocol for a process evaluation. IHC Working Paper 2016. http://www.informedhealthchoices.org/wp-content/uploads/2016/08/IHC-Process-Evaluation-School-resources_final.pdf (accessed May 3, 2017).

14 Austvoll-Dahlgren A, Nsangi A, Semakula D. Interventions and assessment tools addressing key concepts people need to know to appraise claims about treatment effects: a systematic mapping review. *Syst Rev* 2016; **5:** 215.

15 Austvoll-Dahlgren A, Oxman AD, Chalmers I, et al. Key concepts that people need to understand to assess claims about treatment effects. *J Evid Based Med* 2015; **8:** 112–25.

16 Austvoll-Dahlgren A, Semakula D, Nsangi A, et al. The development of the "Claim Evaluation Tools": assessing critical thinking about effects. *BMJ Open* (in press).

17 Austvoll-Dahlgren A, Guttersrud Ø, Semakula D, Nsangi A, Oxman AD. Measuring ability to assess claims about treatment effects: A latent trait analysis of the Claim Evaluation Tools using Rasch modelling. *BMJ Open* (in press).

18 Semakula D, Nsangi A, Guttersrud Ø, et al. Measuring ability to assess claims about treatment effects in English and Luganda: evaluation of multiple-choice questions from the "Claim Evaluation Tools" database using Rasch modelling. IHC Working Paper, 2017. http://www.informedhealthchoices.org/wp-content/uploads/2016/08/Claim-2nd-Rasch-analysis-in-Uganda-2017-03-17.pdf (accessed April 28, 2017).

19 Semakula D, Nsangi A, Oxman AD, et al. Effects of the Informed Health Choices podcast on the ability of parents of primary school children in Uganda to assess claims about treatment effects: a randomised controlled trial. *Lancet* 2017; published online May 21. http://dx.doi.org/10.1016/S0140-6736(17)31225-4.

20 McEwan PJ. Improving learning in primary schools of developing countries: a meta-analysis of randomized experiments. *Rev Educ Res* 2015; **85:** 353–94.

21 Zopluoglu C. A cross-national comparison of intra-class correlation coefficient in educational achievement outcomes. *J Meas Eval Educ Psychol* 2012; **3:** 242–78.

22 Lee DS. Training, wages, and sample selection: estimating sharp bounds on treatment effects. *Rev Econ Stud* 2009; **76:** 1071–102.

23 Snilstveit, B, Stevenson, J, Phillips, D, et al. Interventions for improving learning outcomes and access to education in low- and middle-income countries: a systematic review. London: International Initiative for Impact Evaluation, 2015. http://www.3ieimpact.org/en/evidence/systematic-reviews/details/259/ (accessed Jan 17, 2017).

24 White IR, Thomas J. Standardized mean differences in individually-randomized and cluster randomized trials, with applications to meta-analysis. *Clin Trials* 2005; **2:** 141–51.

25 Seftor N. Raising the bar. *Eval Rev* 2016; published online Sep 7. DOI:10.1177/0193841X16665023.

26 US Department of Education's Institute of Education Sciences. What Works Clearinghouse. Procedures and standards handbook. Version 3.0. Washington, DC: US Department of Education, Institute of Education Sciences. https://ies.ed.gov/ncee/wwc/Docs/referenceresources/wwc_procedures_v3_0_standards_handbook.pdf (accessed Jan 17, 2017).

27 Deke J, Chiang H. The WWC attrition standard: sensitivity to assumptions and opportunities for refining and adapting to new contexts. *Eval Rev* 2016; DOI:10.1177/0193841X16670047.

28 Slavin R, Madden NA. Measures inherent to treatments in program effectiveness reviews. *J Res Educ Eff* 2011; **4:** 370–80.

29 Mpyangu CM, Ochen EA, Onyango EO, Lubaale YAM. Out of school children study in Uganda. Kampala: UN Children's Fund, Uganda Country Office, 2014.

30 Nordheim L, Gundersen MW, Espehaug B, Guttersrud Ø, Flottorp S. Effects of school-based educational interventions for enhancing adolescents' abilities in critical appraisal of health claims: a systematic review. *PLoS One* 2016; **11:** e0161485.

THE LANCET

Supplementary appendix

This appendix formed part of the original submission and has been peer reviewed. We post it as supplied by the authors.

Supplement to: Nsangi A, Semakula D, Oxman AD, et al. Effects of the Informed Health Choices primary school intervention on the ability of children in Uganda to assess the reliability of claims about treatment effects: a cluster-randomised controlled trial. *Lancet* 2017; published online May 21. http://dx.doi.org/10.1016/S0140-6736(17)31226-6.

Effects of using the Informed Health Choices primary school resources on the ability of children in Uganda to assess the reliability of claims about treatment effects: a cluster-randomised trial

Web appendix

Contents

Links to Informed Health Choices resources

Informed Health Choices
Project webpages

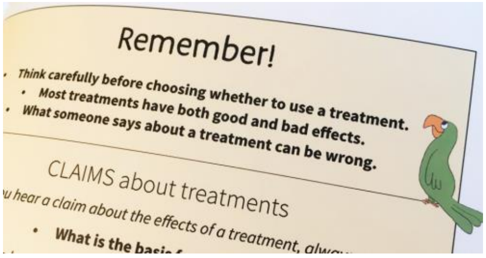
Contact us **if you would like to print these resources or translate them to other languages.**

The Health Choices Book
This textbook for primary school children (10 to 12 year olds) includes a comic story that introduces and explains 12 Key Concepts, instructions for classroom activities, exercises, a list of key messages, and a glossary.

Teachers' Guide
This guide includes lesson plans and other resources to help teachers using The Health Choices Book.

Exercise Book
This includes the exercises from The Health Choices Book.

Activity Cards
These cards are for use in Lesson 7 of the The Health Choices Book. The activity is created to demonstrate how comparisons with few people can be misleading.



Checklist

This poster with the key messages from The Health Choices Book is a checklist for applying the 12 Key Concepts and a reminder of the most important messages in the book.

Think Carefully about Treatments
Song credits: Informed Health Choices team & Swangs Avenue

The lyrics of this song are another reminder of the key messages in the book.

Claim Evaluation Tools

The Claim Evaluation Tools consist of multiple-choice questions that assess an individual's ability to apply the Key Concepts.

Spiral curriculum

We will connect the resources that we are developing in a spiral curriculum. This will serve as a map, outlining where learners should begin and how they should progress.

Nsangi A, Semakula D, Oxman AD, et al. Does the use of the Informed Healthcare Choices (IHC) primary school resources improve the ability of grade-5 children in Uganda to assess the trustworthiness of claims about the effects of treatments: protocol for a cluster-randomised trial. Trials, in press.

Nsangi A, Semakula D, Rosenbaum, et al. Development of the Informed Health Choices resources to teach primary school children to assess claims about treatment effects in four countries. IHC Working Paper, 2017.

Nsangi A, Semakula D, Glenton C, et al. Resources to teach primary school children in low-income countries to assess claims about treatment effects: protocol for a process evaluation. IHC Working Paper 2016.

Davies A, Gerrity M, Nordheim L, et al. Measuring ability to assess claims about treatment effects: establishment of a standard for passing and mastery. IHC Working Paper 2017.

Nsangi A, Oxman AD, Sewankambo NK. Teaching children in low income countries to assess claims about treatment effects; a prioritisation of key concepts. J Evid Based Med 2015; 8: 173-80.

Web table 1: Intended behaviours

Think about an illness that you might get. Imagine someone claiming (saying) that a particular treatment might help you get better.

	How likely are you to find out what the claim was based on (for example by asking the person making the claim)?		How likely are you to find out if the claim was based on a research study comparing the treatment to no treatment (a fair comparison)?		How likely are you to say "yes" if you are asked to participate in a research study comparing two treatments for your illness (a fair comparison)?	
	Control schools N=4430	**Intervention schools** N=5753	**Control schools** N=4430	**Intervention schools** N=5753	**Control schools** N=4430	**Intervention schools** N=5753
Missing	211 (4.8%)	189 (3.3%)	236 (5.3%)	211 (3.7%)	198 (4.5%)	187 (3.3%)
I don't know	1015 (23.0%)	775 (13.4%)	899 (20.3%)	884 (15.3%)	733 (16.6%)	754 (13.1%)
Very unlikely	395 (8.9%)	589 (10.2%)	610 (13.8%)	699 (12.2%)	343 (7.7%)	372 (6.5%)
Unlikely	369 (8.3%)	469 (8.2%)	718 (16.2%)	845 (14.7%)	389 (8.8%)	520 (9.0%)
Likely	1517 (34.2%)	2018 (35.1%)	1025 (23.1%)	1593 (27.7%)	1552 (35.0%)	2228 (38.7%)
Very likely	923 (20.8%)	1713 (29.8%)	942 (21.3%)	1521 (26.4%)	1215 (27.4%)	1692 (29.4%)
Likely or very likely*	**2440 (55.1%)**	**3731 (64.9%)**	**1967 (44.4%)**	**3114 (54.1%)**	**2163 (48.8%)**	**3201 (55.6%)**
Odds ratio†	1.56 (95% CI 1.29 to 1.88)		1.54 (95% CI 1.29 to 1.84)		1.37 (95% CI 1.16 to 1.62)	
Adjusted Difference†	**10.6%** (95% CI 6.2% to 14.7%)		**10.8%** (95% CI 6.3% to 15.1%)		**7.8%** (95% CI 3.7% to 11.9%)	

* Missing values and don't know are pooled with unlikely and very unlikely.

† The difference is an adjusted difference, based on mixed models with a random effects term for the clusters and the stratification variables modelled as fixed effects, using logistic regression. P < 0.001 for all three comparisons. The odds ratios from the logistic regressions have been converted to differences based using the control schools as the reference.

Web table 2: Self-efficacy

How difficult or easy would you find each of these actions to be?

	Assessing whether a claim about a treatment is based on a research study comparing treatments (a fair comparison)		**Assessing where I can find information about treatments that is based on research studies comparing treatments (fair comparisons)**		**Assessing how sure I can be about the results of a research study comparing treatments (the trustworthiness of the results)**		**Assessing if the results of a research study comparing treatments are likely to be relevant to me**	
	Control schools N=4430	Intervention schools N=5753	Control schools N=4430	Intervention schools N=5753	Control schools N=4430	Intervention schools N=5753	Control schools N=4430	Intervention schools N=5753
Missing	190 (4.3%)	171 (3.0%)	208 (4.7%)	202 (3.5%)	221 (5.0%)	212 (3.7%)	194 (4.4%)	194 (3.4%)
I don't know	688 (15.5%)	589 (10.2%)	584 (13.2%)	618 (10.8%)	823 (18.6%)	1084 (18.8%)	767 (17.3%)	1019 (17.7%)
Very difficult	595 (13.4%)	617 (10.7%)	545 (12.3%)	675 (11.7%)	696 (15.7%)	912 (15.9%)	699 (15.8%)	853 (14.8%)
Difficult	1071 (24.2%)	1132 (19.7%)	855 (19.3%)	1189 (20.7%)	913 (20.6%)	1433 (24.9%)	768 (17.3%)	690 (16.7%)
Easy	1332 (30.1%)	2154 (37.4%)	1387 (31.3%)	1803 (31.3%)	1058 (23.9%)	1183 (20.6%)	1039 (23.5%)	1430 (24.9%)
Very easy	554 (12.5%)	1090 (18.9%)	851 (19.2%)	1266 (22.0%)	719 (16.2%)	929 (16.1%)	963 (21.7%)	1297 (22.5%)
Easy or very easy*	**1886 (42.6%)**	**3244 (56.4%)**	**3069 (53.3%)**	**2238 (50.5%)**	**1777 (40.1%)**	**2112 (36.7%)**	**2002 (45.2%)**	**2727 (47.4%)**
Odds ratio†	1.83 (95% CI 1.55 to 2.16)		1.13 (95% CI 0.96 to 1.33)		0.84 (95% CI 0.73 to 0.96)		1.08 (95% CI 0.93 to 1.25)	
Adjusted difference†	**15.0%** (95% CI 10.9% to 19.0%)		**3.0%** (95% CI -1.0% to 7.0%)		**-4.1%** (95% CI -1.0% to -7.3%)		**1.9%** (95% CI -1.8% to 5.6%)	

* Missing values and don't know are pooled with difficult and very difficult.

† The difference is an adjusted difference, based on mixed models with a random effects term for the clusters and the stratification variables modelled as fixed effects, using logistic regression. The odds ratios for the comparison in the last column from the logistic regressions have been converted to differences based on the intervention school proportions and the odds ratios calculated using the control schools as the reference.

Web table 3: Attitudes

	School		Science	
	How much do you like school?		How much do you like science as a subject?	
	Control schools N=4430	Intervention schools N=5753	Control schools N=4430	Intervention schools N=5753
Missing	55 (1.2%)	40 (0.7%)	100 (2.3%)	121 (2.1%)
Dislike a lot	119 (2.7%)	181 (3.1%)	123 (2.8%)	191 (3.3%)
Dislike	85 (1.9%)	130 (2.3%)	120 (2.7%)	156 (2.7%)
Like	788 (17.8%)	955 (16.6%)	1118 (25.2%)	1541 (26.8%)
Like a lot	3383 (76.4%)	4447 (77.3%)	2969 (67.0%)	3744 (65.1%)
Like or like a lot*	**4171 (94.2%)**	**5402 (93.9%)**	**4087 (92.3%)**	**5285 (91.9%)**
Odds ratio†	0.96 (95% CI 0.69 to 1.34)		1.08 (95% CI 0.78 to 1.33)	
Adjusted difference†	**-0.2%** (95% CI -2.4% to 1.4%)		**-0.1%** (95% CI -2.7% to 1.8%)	
	How helpful to you is what you are learning in school?		How helpful to you is what you are learning in science?	
	Control schools N=4430	Intervention schools N=5753	Control schools N=4430	Intervention schools N=5753
Missing	66 (1.5%)	68 (1.2%)	104 (2.3%)	125 (2.2%)
Very unhelpful	202 (4.6%)	265 (4.6%)	178 (4.0%)	273 (4.7%)
Unhelpful	119 (2.7%)	155 (2.7%)	121 (2.7%)	186 (3.2%)
Helpful	588 (13.3%)	847 (14.7%)	594 (13.4%)	991 (17.2%)
Very helpful	3455 (77.9%)	4418 (76.8%)	3433 (77.6%)	4178 (72.7%)
Helpful or very helpful*	**4043 (91.3%)**	**5265 (91.5%)**	**4207 (95.0%)**	**5169 (89.8%)**
Odds ratio†	0.98 (95% CI 0.72 to 1.33)		0.96 (95% CI 0.70 to 1.32)	
Adjusted difference†	**0.6%** (95% CI -2.2% to 2.0%)		**-0.2%** (95% CI - 2% to 1.2%)	

* Missing values are pooled with dislike a lot and dislike, or very unhelpful and unhelpful.

† The difference is an adjusted difference, based on mixed models with a random effects term for the clusters and the stratification variables modelled as fixed effects, using logistic regression. The odds ratios for the comparison in the last column from the logistic regressions (0.96, 95% CI 0.69 to 1.34; 1.08, 95% CI 0.78 to 1.33; 0.98, 95% CI 0.72 to 1.33; and 0.96, 95% CI 0.70 to 1.32) have been converted to differences based on the intervention school proportions and the odds ratios calculated using the control schools as the reference.

Web table 4: Intervention school children's views of *The Health Choices Book*

How much did you like what you learned as part of the lessons with "The Health Choices Book"?		How helpful to you is what you learned as part of the lessons with "The Health Choices Book"?		How easy or difficult to understand did you find the lessons with "The Health Choices Book"?		Do you trust what you learned as part of the lessons with "The Health Choices Book"?	
Intervention schools N=5753							
Missing	125 (2.2%)	Missing	138 (2.4%)	Missing	52 (0.9%)	Missing	54 (0.9%)
Dislike a lot	130 (2.3%)	Very unhelpful	282 (4.9%)	Very difficult	342 (5.9%)	Very little trust	299 (5.2%)
Dislike	104 (1.8%)	Unhelpful	185 (3.2%)	Difficult	495 (8.6%)	Little trust	361 (6.3%)
Like	918 (16.0%)	Helpful	987 (17.2%)	Easy	1853 (32.2%)	Trust it	1112 (19.3%)
Like a lot	4476 (77.7%)	Very helpful	4161 (72.3%)	Very easy	3011 (52.4%)	Trust it very much	3927 (68.3%)
Like or like a lot*	**5394 (93.8%)**	**Helpful or very helpful***	**5148 (89.5%)**	**Easy or very easy***	**4864 (84.5%)**	**Trust it or trust it very much**	**5039 (87.6%)**

* Missing values are pooled with dislike a lot and dislike, very unhelpful and unhelpful, very difficult and difficult, or very little trust and little trust.

Web table 5: Subgroup analysis - reading skills

	Control schools	Intervention schools	Adjusted difference†	Odds ratio	ICC
Mean score, %					
Lacking basic reading skills (N=4363)	N children = 2139 Mean score: 39.1% (SD 13.8%)	N children = 2224 Mean score: 54.0% (SD 17.2%)	**Mean difference: 15.7%** (95% CI 13.2% to 18.3%)		0.17
Basic reading skills (N=3662)	N children = 1507 Mean score: 46.3% (SD 15.5%)	N children = 2155 Mean score: 65.6% (SD 17.7%)	**Mean difference: 20.7%** (95% CI 18.2% to 23.3%)		0.14
Advanced reading skills (N=2098)	N children = 766 Mean score: 47.8% (SD 15.7%)	N children = 1332 Mean score: 71.0% (SD 17.3%)	**Mean difference: 24.0%** (95% CI 20.9% to 27.1%)		0.17
Passing score (≥ 13 out of 24 correct answers)					
Lacking basic reading skills	N children = 2139 17.2% of children N=368	N children = 2224 51.4% of children N=1143	**39.8% more children** (95% CI 31.3% to 47.9%)	6.38 (95% CI 4.53 to 8.99)	0.16
Basic reading skills	N children = 1507 34.7% of children N= 523	N children = 2155 76.6% of children N= 1651	**48.8% more children** (95% CI 43.5% to 52.9%)	9.49 (95% CI 6.75 to 13.34)	0.15
Advanced reading skills	N children = 766 38.0% of children N=291	N children = 1332 85.7% of children N=1142	**52.5% more children** (95% CI 48.1% to 55.6%)	15.56 (95% CI 10.12 – 23.93)	0.16
Mastery score (≥ 20 out of 24 correct answers)					
Lacking basic reading skills	N children = 2139 0.3% of children N=7	N children = 2224 6.9% of children N=154	**7.0% more children** (95% CI 2.9% to 15.6%)	26.26 (95% CI 10.98 to 62.79)	0.20
Basic reading skills	N children = 1507 1.1% of children N=16	N children = 2155 22.0% of children N=475	**26.5% more children** (95% CI 15.6% to 48.8%)	34.30 (95% CI 17.98 to 65.44)	0.15
Advanced reading skills	N children = 766 2.0% of children N=15	N children = 1332 32.6% of children N=434	**38.2% more children** (95% CI 23.3% to 55.7%)	30.34 (95% CI 14.91 to 61.72)	0.18

* Because reading skills were measured after the intervention, we have not reported a test of interaction here (see Appendix 3).

† The adjusted difference is based on mixed models with a random effects term for the clusters and the stratification variables modelled as fixed effects, using logistic regression for dichotomous outcomes and linear regression for continuous outcomes. The odds ratios from the logistic regressions for passing scores and mastery scores have been converted to differences based on the control school proportions and the odds ratios shown here.

Web table 6: Costs

	Unit cost	N	Cost	Years use*	Cost per Year†
Children's book	$5.61	5753	$32,274	5	$7455
Teachers' guide	$49.38	85	$4197	5	$969
Exercise book	$1.30	5753	$7479	1	$7479
Test	$0.54	5753	$3107	1	$3107
Activity cards	$8.94	85	$760	5	$176
Poster	$4.35	85	$370	5	$85
Delivery of materials	$30.77	85	$2615	1	$2615
Teacher workshops‡	$64.42	85	$5476	5	$1265
Total			$64,370		**$23,151**
Cost per school (not including teachers' time)					$385.84
Cost per child (not including teachers' time)					$4.02
Teachers' time§					
Teachers' classroom time (months)	$125	0.17	$1811	1	$1811
Teachers' training time (months)	$125	0.09	$966	5	$223
Total teachers' time			$5385		**$2,034**
Cost per school (including teachers' time)					$419.75
Cost per child (including teachers' time)					$4.38

* We assumed that the teaching materials, apart from the exercise book and the test, would be used over a five-year period and the training workshops for the teachers would not need to be repeated during this time.

† Based on an interest rate of 5% and an annualization factor of 0.2310

‡ Based on the estimated cost of four workshops, including travel costs, room rental, food, materials, and trainer time (including trainer preparation)

§ No additional resources were introduced to pay for the teachers' time. However, this is an opportunity cost, since the teachers' time that was used for the Informed Health Choices lessons could not be used for something else.

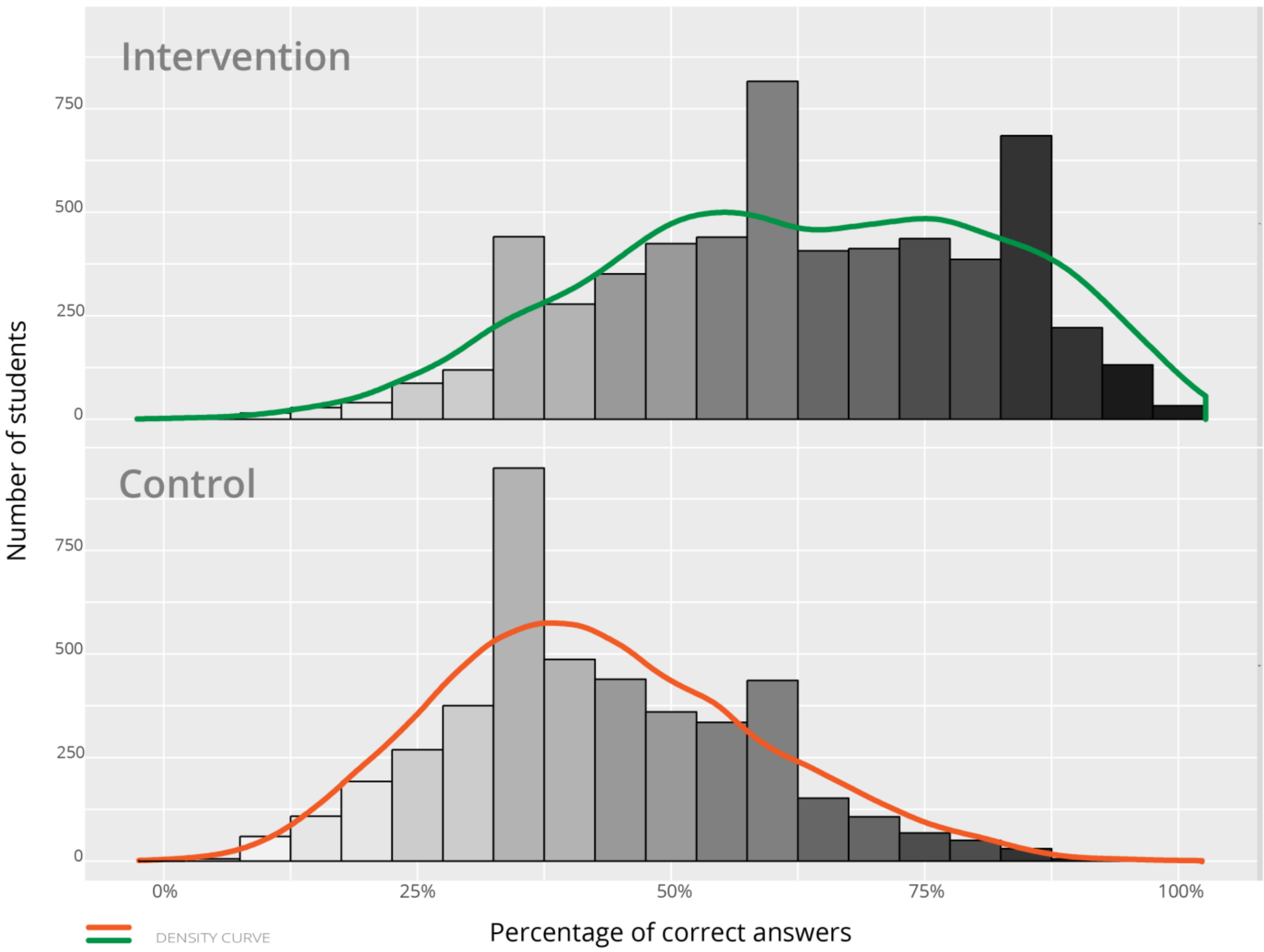
Web table 7: Potential adverse and corresponding beneficial effects

Potential adverse effects	*Corresponding beneficial effects*
Conflict between children and teachers due to children challenging teachers' autority	**More open and engaging discussion** of the basis of diverse claims and beliefs
Conflict between children and parents due to children challenging parents' authority	**Better understanding between children and parents** due to children conversing with their parents about what they are learning and parents feeling more engaged with what their children are learning + engagement of parents in discussions of health issues
Distrust of health professionals or **conflict between children and health professionals**	**Appropriate questioning of health professionals, better understanding and better healthcare**
Conflict due to undermining of religious beliefs	**Engagement of children and others in discussion** about religious beliefs and science
Shortened enjoyment of the innocence of childhood	**Increased enjoyment of school and childhood**
Nihilism or cynicism	**Healthy scepticism and appreciation of science**

Web figure 1: The four districts included in the trial

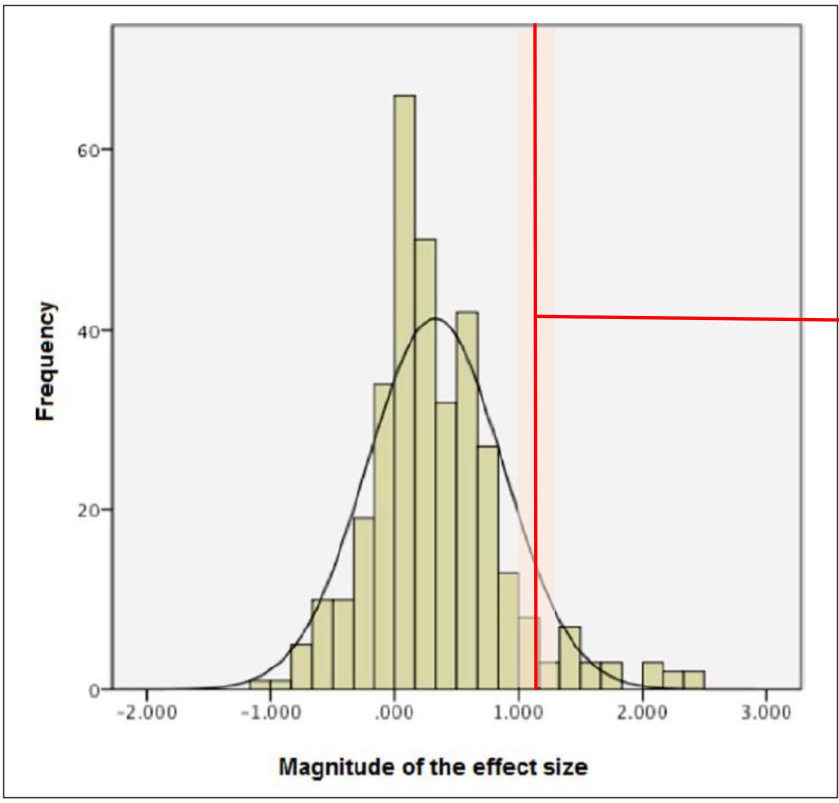
The four districts included in the trial were Kampala, Luweero, Mukono, and Wakiso.

Web figure 2: Distribution of scores and density curves



Web figure 3: Comparison of other strategies for teaching critical thinking

Effect sizes (Hedges' g) for strategies for teaching critical thinking*



Informed Health Choices primary school intervention
1.16 (95% CI 1.00 to 1.32)

FIGURE 1. *Distribution of unweighted effect sizes for generic critical thinking skills (k = 341, $\bar{X}$ = 0.33, SD = 0.55).*

*From Abrami PC, Bernard RM, Borokhovski E, Waddington DI, Wade CA, Persson T. Strategies for teaching students to think critically a meta-analysis. Rev Educ Res 2015; 85: 275-314.

Appendix 1: GREET 2015 checklist

GREET 2015 checklist,[1] Based upon the TIDieR guidance[2]

BRIEF NAME
1. INTERVENTION: We conducted a cluster randomised trial to compare use of the Informed Health Choices (IHC) primary school resources (intervention) to routine teaching (control) in primary schools in Uganda.
WHY - this educational process
2. THEORY: The IHC primary school resources were developed by the investigators between 2013 and 2015 employing user-centred design methods. This included idea generation and prototyping, piloting with observation, user-testing with teachers and learners, and teacher's network feedback in Uganda. Piloting and user-testing were also carried out in Kenya, Rwanda, and Norway. The aim of the design process was to ensure that teachers and children found the learning resources to be engaging and useful. 3. LEARNING OBJECTIVES: The objectives were for children to understand and be able to apply key concepts[3] for assessing claims about the effects of treatments (any action intended to improve the health of individuals or communities) and to making informed health choices. 4. EBP CONTENT: The learning resources focused on 12 key concepts for appraising claims and evidence about treatment effects and applying appraised evidence to personal choices.
WHAT
5. MATERIALS: Teachers in the intervention schools attended a two-day introductory workshop. We gave them a teachers' guide prior to participation in the workshops. We gave the intervention schools textbooks and exercise books for the children, activity cards for one of the activities in the textbook and teachers' guide, and a poster with a checklist summarising the 12 key concepts covered by the book. We gave them a song (Think Carefully about Treatments) with lyrics that are another reminder on MP3 players for the final lesson. The textbooks included a story told in a comic book format, instructions for classroom interactive activities, exercises, the checklist, a glossary, and a gameboard on the back of the book for another classroom activity. The textbook included nine chapters with exercises and an activity for each. Two of the investigators took the teachers through each chapter during the introductory workshops. All the materials can be accessed on the IHC website http://www.informedhealthchoices.org/primary-school-resources/. 6. EDUCATIONAL STRATEGIES: We designed the materials to be used interactively in the classroom by reading each chapter aloud, doing the activity and, if time allowed, giving the children time to do the exercises. The exercise books could be taken home, if there was not time to do the exercises in the classroom. Educational strategies that we used included repetition of key messages, extensive examples familiar to the children, visual presentation (comic format), messages embedded in a narrative, defining new vocabulary where it is introduced and translating words to Luganda and Swahili, activities that require interaction between students, a highly-structured timetable for teachers, additional explanations and examples for teachers. 7. INCENTIVES: The head teacher in each participating school selected the teachers. The teachers were reimbursed for travel costs for the introductory workshop and received meals and refreshment. They were not paid for participating in the workshop and there were no financial incentives for the schools, head teachers, teachers, or children. The evaluation administered at the end of the school term did not count towards the children's school marks or assessment of the teachers or schools.
WHO PROVIDED
8. INSTRUCTORS: The teachers were year-5 teachers. Most (80%) were science teachers. Only 12% had a university degree.

HOW
9. DELIVERY: The nine lessons were delivered in the classroom. The average number of children in each class was 72.
WHERE
10. ENVIRONMENT: The primary schools were in the Central region of Uganda. Most (68%) were in an urban area. Half were public and half were private schools. All the schools were poorly resourced with respect to space (crowded classrooms with too few benches), equipment (little or no access to computers or other electronic equipment), and supplies.
WHEN and HOW MUCH
11. SCHEDULE: There was one lesson for each chapter. The nine lessons were taught in a single school term lasting 10 to 12 weeks. Each school decided how to fit the lessons into the term, for the most part they taught one lesson per week. 12. A double period (80 minutes) was recommended for each lesson, so that the total amount of class time was nine double periods (12 hours). The amount of time that the children spent on the lessons outside of class varied, but for the most part was little if any, since most teachers did not allow the children to take the textbooks home and most of the exercises were done during the double periods, in most schools. We suggested that the teachers should spend about 20 minutes preparing for each lesson.
PLANNED CHANGES
13. The teachers' guide included options for the teachers, such as different ways of reading the text aloud and different ways of marking the exercises and giving the children feedback.
UNPLANNED CHANGES
14. Each teacher was observed by the research team for one lesson, but no feedback was given to the teachers. Some of the teachers improvised their own activities.
HOW WELL
15. ATTENDANCE: Attendance varied. The intervention did not include any strategies for improving attendance. 16. Fidelity will be reported in a process evaluation. Teachers completed an evaluation form for each lesson, the research team observed each teacher teaching a lesson, and we interviewed head teachers, teachers and children in six schools. Analysis of these data has not been completed. 17. All the classes completed all nine lessons, but not all the children attended all nine lessons and some of the teachers did not use two full periods for each lesson.

1. Phillips AC, Lewis LK, McEvoy MP, et al. Development and validation of the guideline for reporting evidence-based practice educational interventions and teaching (GREET). BMC medical education. 2016; 16: 237.

2. Hoffmann TC, Glasziou PP, Boutron I, et al. Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. BMJ 2014; 348: g1687.

3. Austvoll-Dahlgren A, Oxman AD, Chalmers I, Nsangi A, Glenton C, Lewin S, et al. Key concepts that people need to understand to assess claims about treatment effects. Journal of Evid Based Med 2015; 8: 112-25.

Appendix 2. Claim test taken at the end of the term

This questionnaire includes multiple-choice questions about treatment claims. Please answer all questions to the best of your ability.

The questionnaire includes some words that may be unfamiliar to you:

A **TREATMENT** is anything done to care for yourself, so you stay well or, if you are sick or injured, so you get better and not worse. For example, wearing glasses (to see better).
IN LUGANDA: OBUJJANJABI

A **TREATMENT CLAIM** is something someone says about whether a treatment causes something to happen or to change. A claim can be true or can be false. For example, that wearing glasses makes you see better.
IN LUGANDA: EKINTU EKYOGERWAYOGERWA KUBY'OBUJJANJABI

A **RESEARCH STUDY** is a way to answer a question by carefully collecting information. For example, a study might be done to answer the question: Does wearing glasses make people see better?
IN LUGANDA: OKUNOONYEREZA OKWEKINNASAYANSI

RESULTS of a study are what the study found. For example, whether people who wear glasses could see better.
IN LUGANDA: EKIVAAMU MUKUNOONYEREZA

Part 1. Questions about you

1.1 **How old are you?** __________

1.2 **Are you a:**

□ Girl

□ Boy

Part 2. Questions about claims

Instructions: Read the passage on every question then answer the question below the passage using one of the provided answers. For each question, choose what you think is the best answer and write the letter for that answer in the box provided.

Example

A teacher says that the children in his school run faster than the children going to school in another village.

Question: **How can the teacher be sure about this?**

Options:

A) He should ask a teacher at the other school

B) He should arrange for a running contest between the two schools

C) He should ask the children in his school what they think

D) He should ask the children in the other school what they think

Answer: B

Start

2. A doctor did a research study to find out if drinking tea keeps people from getting sick. He tossed a coin to decide who should get the tea and who should not. People who got tea went to the doctor's office every day to drink their tea. At the end of the study, people who got the tea were less likely to be sick than those who got no tea.

Based on the text above, please answer the following questions:

2.1 Who went to the doctor's office every day?

Options:

A) People who did not get tea

B) People who got tea

C) Everyone

D) People who got sick

Answer: ☐

2.2 How did the doctor decide who should get tea?

Options:

A) By tossing a coin

B) By asking people what they would like

C) The doctor gave tea to those who were more likely to be sick

D) The doctor asked people who came to his office

Answer: ☐

3. A doctor did a research study to find out if drinking tea keeps people from getting sick. He tossed a coin to decide who should get the tea and who should not. People who got tea went to the doctor's office every day to drink their tea. At the end of the study, people who got the tea were less likely to be sick than those who got no tea.

Based on the text above, please answer the following questions:

3.1 What was the treatment?

Options:

A) Tea

B) Sleep

C) The study

D) The doctor

Answer: ☐

3.2 What was the result of the study?

Options:

A) Drinking tea can help people from getting sick

B) Doctors should toss coins when doing studies

C) People should go to the doctor if they are sick

D) Not drinking tea can help people from getting sick

Answer: ☐

4. Annette sees an advert on TV for a new soap which the makers say protects people from getting skin rashes. Annette thinks that this soap must be better than other soaps for protecting her skin.

Question: **Is Annette right?**

Options:

A) No, the soap may be newer, but that does not mean that it is better than other soaps

B) Yes, the new soap is probably better than most other soaps because it is newer

C) Yes, the new soap is probably better than most other soaps because a well-known company makes it

Answer: ☐

5. Regina has an illness that makes it difficult to breathe. She hears on the radio about a medicine that has helped many people for their breathing problems.

Question: **How sure can Regina be that the medicine does not have any harms?**

Options:

A) It is not possible to say. However, medicines are rarely harmful

B) Not very sure, because all medicines may harm people as well as help them

C) Very sure, since the medicine has helped many people, it is unlikely that it also harms people

Answer: ☐

6. John has a skin rash on his leg. A shop sells several creams to treat skin rashes. John chooses a cream from a well-known company, even though it is more expensive than the other creams. John thinks the cream is more likely to heal his rash than the other creams because it is more expensive.

Question: **Is John right?**

Options:

A) No, just because the cream is expensive does not mean that it will work better than other creams

B) It is not possible to say. However, expensive creams are likely to be better because the companies spend more time making them

C) No, the cream is probably not as good as the other creams. Well-known companies are usually better at advertising

D) Yes, the company is well-known for a reason, so it is more likely to be better than creams sold by lesser-known companies

Answer:

7. Two companies make two different medicines for treating stomach pain. Each of them says that their medicine is the better one.

Question: **How can you know which of the two medicines is better for stomach pain?**

Options:

A) It is not possible to say. The companies may just say their medicine is best because they want to make money

B) I would rely on the best-known company; it is more likely to have the best medicine

C) I cannot trust either of the companies. They are probably both wrong

Answer:

8. Dr. Kato and Dr. Semakula disagree about which medicine for stomach pain is best. Dr. Kato says his opinion is right because he has worked as a doctor for a longer time than Dr. Semakula.

Question: **Is Dr. Kato right?**

Options:

A) Yes, because Dr. Kato has worked for a long time, he has more experience than Dr. Semakula

B) Yes, because Dr. Kato has worked for a long time, he must be basing what he says on studies comparing the medicines

C) No, just because Dr. Kato has worked as a doctor for a longer time does not mean that he is basing what he says on studies that compare medicines for stomach pain

Answer: ☐

9. Habibah has pain in her ear, and she asks her brother Hassan what to do about it. He says that once, when he had a pain like that, he rinsed his ear with hot water. The next day, his ear pain was gone. Based on his experience, he says rinsing with hot water is helpful for ear pain.

Question: **Do you agree with Hassan?**

Options:

A) Yes. Because this is Hassan's experience, it is likely to be true

B) No, Hassan's experience is not enough to be sure

C) Yes, Hassan rinsed his ear with hot water and the next day his ear pain was gone

Answer: ☐

10. Sarah has an illness. There is a medicine for it, but she is unsure if she should try it. A research study comparing the medicine with no medicine found that the medicine was helpful but also that it could be harmful. Three of Sarah's friends are giving her advice about what to do.

Question: **Which advice given to her by her friends is the best advice?**

Options:

A) She should only take the medicine if many people have tried the medicine before

B) She should only take the medicine if she thinks it will help her more than it will harm her

C) If Sarah has enough money to buy the medicine, it could not hurt to try it

Answer:

11. Dr. Acheng is an expert on treating headaches. A news reporter interviews Dr. Acheng about a new medicine. Dr. Acheng says that, in her personal experience, the new medicine is good for treating headaches.

Question: **How sure can we be that Dr. Acheng right?**

Options:

A) It is not possible to say. It depends on how long Dr. Acheng has been an expert on treating headaches

B) Not very sure. Even though Dr. Acheng is an expert, the new medicine still needs to be compared in studies with other treatments

C) Very sure. Dr. Acheng is an expert, so she knows if the new medicine is good or not based on her experience

D) Very sure. Dr. Acheng would not be interviewed by a news reporter if her advice was not good

Answer: ☐

12. Edith has a stomach pain. Edith's mother says that fruit juice is a good treatment for stomach pain. She learnt about this treatment from Edith's grandmother. Over many years, other families she knows have also used fruit juice to treat stomach pain.

Question: **Based on this, how sure can we be that fruit juice is a good treatment for stomach pain?**

Options:

A) Not very sure. Even though people have used fruit juice over many years, that does not mean that it helps stomach pain

B) Very sure. If it has worked for Edith's mother and other people who have tried it, it will probably work for her too

C) Not very sure. Edith should ask more families if they use fruit juice to treat stomach pain

Answer: $\square$

13. At David's school, some students have poor parents. The students with poor parents drink less fruit juice than the children of other parents. The students with poor parents are also more often sick. Based on this link, David thinks that people who drink fruit juice, are less likely to get sick.

Question: **Is David correct?**

Options:

A) It is not possible to say, it depends on whether or not Peter has poor parents

B) Yes, students with poor parents do not drink fruit juice and are more often sick

C) Yes, the juice is the only possible reason why the students with the poor parents are more often sick

D) It is not possible to say. There could be other reasons why students with poor parents are more often sick

Answer: ☐

14. In a research study done by John, four people were told to do exercises every day for a month, and four people were told to eat bananas every day. At the end of the month, the people who ate bananas had more strength than those who did exercises. Based on his study, John advises his friend Mildred to eat bananas.

Question: **Mildred says that we cannot be sure about the results of John's study. Why?**

Options:

A) Because the study included so few people, the differences in strength could have happened by chance, and not because of the bananas

B) Because John should have included fewer people in his study so that he could have followed them more closely

C) Because four people is not enough, all people taking part in the study should have been told to eat the bananas

Answer: ☐

15. A new fruit drink is said to make people feel strong. Fred wanted to know if this is true, and decided to do a research study comparing people who got the new fruit drink and people who drank just water.

People in the study knew if they got the new drink or water, and Fred told them that the new fruit drink was likely to make people stronger. At the end of the study, Fred was right and those who drank the new fruit drink said they felt stronger.

Question: **Why can't we be sure about the results of Fred's study?**

Options:

A) Because all people taking part in the study should have been given the new fruit drink

B) Because people knew if they got the new fruit drink, and knowing this may have influenced how they felt

C) Because Fred should have told both groups that they could expect to feel stronger

Answer:

16. Harriet is worried about getting sick. She hears about a new research study on the radio that compared a new medicine to an old medicine. Fewer people who took the old medicine got sick compared to the people who took the new medicine.

Question: **How sure can Harriet be that the old medicine is better than the new medicine?**

Options:

A) Less sure, because Harriet needs to know the results of all other studies comparing the new medicine with the old medicine

B) More sure, because she heard about the study on the radio

C) Less sure, unless she finds another study with the same results

D) More sure, because this is a new study

Answer:

17. Doctors studied people with stomach pain before and after they took a new medicine. After taking the new medicine, many people felt less pain.

Question: **Can we be sure that the new medicine is good for treating stomach pain?**

Options:

A) No, taking the new medicine should have been compared either with not taking the medicine, or with taking an older medicine

B) Yes, people were asked how much pain they felt before and after they took the new medicine

C) Yes, the study was done by doctors

Answer: ☐

18. In a research study, doctors compared two treatments for knee pain, a new and an old treatment. People were able to choose which treatment they got. Most young people chose the new treatment. At the end of the study, people who chose the new treatment had less pain.

Question: **How sure can you be that the new treatment is better for treating pain than the old treatment?**

Options:

A) Less sure, because people taking the new and old treatment were not similar

B) Less sure, because all people taking part in the study should have got both treatments

C) Less sure, because older people did not like the new treatment

Answer:

19. Judith wants smoother skin. The younger girls in her school have smoother skin than the older girls. Judith thinks this is because the younger girls use cream on their skin to make the skin smoother.

Question: **Based on this link between using cream and smooth skin, is Judith correct?**

Options:

A) It is not possible to say. It depends on how many younger and older girls there are

B) It is not possible to say. There might be other differences between the younger and older girls

C) Yes, because the younger girls use cream on their skin and they have smoother skin

D) No, Judith should try using the cream herself to see if it works for her

Answer: ☐

20. Dr. Wasswa has done a research study giving a new medicine to people who were vomiting. Some of the people stopped vomiting after they got the new medicine. Dr. Wasswa says that this means that the medicine works.

Question: **Is Dr. Wasswa right?**

Options:

A) No. The people who used the medicine were not compared with similar people who did not use the medicine

B) Yes, some of the people stopped vomiting

C) No, since not all of the people stopped vomiting

Answer: ☐

Instructions: Read the passage at the top of the box. Then read the text in each row and choose what you think is the best answer by making a tick ✓ in one of the two boxes. There should be only one tick in each row.

21. When you are sick, sometimes people say that something - a treatment - is good for you. Below you will find different things people say about such treatments. **Do you agree or disagree with each of the following things being said?**		
For each thing being said below, use a tick ✓ to mark whether you "agree" or "disagree".		
Things being said:	**Agree**	**Disagree**
21.1 Peter says that if a treatment works for one person, the treatment will help others too	☐	☐
21.2 Alice says that if some people try the treatment and feel better, this means that the treatment helps	☐	☐
21.3 Habibah says that, just because many people are using the treatment, this does not mean that it helps	☐	☐
21.4 Julie says that companies sometimes say that the treatment they make is best just to make money	☐	☐

22. A doctor wanted to know if a new medicine for treating headaches is better than an older medicine. The doctor did a research study, comparing the two medicines.

Which of the actions would help us be more sure about the results?

For each action below, use a tick ✓ to mark whether you think the action would help us be "more sure" or "less sure" about the results of the study.

Actions:	**More sure**	**Less sure**
22.1 The doctor should use chance (like tossing a coin) to decide which people should be given the new and which the old medicine	☐	☐
22.2 People should not know which medicine they get (the new medicine or the old medicine) until the end of the study	☐	☐
22.3 The doctor should include only a small number of people in the study	☐	☐

23. To know if a treatment helps you, the treatment should be compared in research studies to other treatments (fair comparisons). Below you will find different things people say about such studies. **Do you agree or disagree with each of the following things being said?**		
For each thing being said below, use a tick ✓ to mark whether you "agree" or "disagree".		
Things being said:	**Agree**	**Disagree**
23.1 Julie says that, if a treatment has been compared in a study to another treatment, you don't have to look for more studies	☐	☐
23.2 Margaret says that the results of a study should be used to decide if a treatment is more helpful than harmful	☐	☐

Part 3. Questions about your views

Instructions: For the following questions, there are no right or wrong answers. Read the passage at the top of the box. Then read the text in each row and choose what you think is the best answer by making a tick ✓ in one of the five boxes. There should be only one tick in each row

24. Think about an illness that you might get. Imagine someone claiming (saying) that a particular treatment might help you get better.

How likely are you to do each of the following actions?

(Mark with a tick ✓ *in the box; one check for each row.)*

Actions:	**Very unlikely**	**Unlikely**	**Likely**	**Very likely**	**I don't know**
24.1 Find out what the claim was based on (for example by asking the person making the claim)	☐	☐	☐	☐	☐
24.2 Find out if the claim was based on a research study comparing the treatment to no treatment (a fair comparison)	☐	☐	☐	☐	☐

25. Below are some actions. Please read each one carefully and give the answer that comes closest to **how difficult or easy you find each of the actions to be:**

(Mark with a tick ✓ in the box; one check for each row.)

Actions:	**Very difficult**	**Difficult**	**Easy**	**Very easy**	**I don't know**
25.1 Assessing whether a claim about a treatment is based on a research study comparing treatments (a fair comparison)	☐	☐	☐	☐	☐
25.2 Assessing where I can find information about treatments that is based on research studies comparing treatments (fair comparisons)	☐	☐	☐	☐	☐
25.3 Assessing how sure I can be about the results of a research study comparing treatments (the trustworthiness of the results)	☐	☐	☐	☐	☐
25.4 Assessing if the results of a research study comparing treatments are likely to be relevant to me	☐	☐	☐	☐	☐

26. Think about an illness that you might get. **How likely are you to say "yes" if you are asked to participate in a research study comparing two treatments for your illness (a fair comparison)?**

(Mark with a tick ✓ in one box)

Very unlikely	Unlikely	Likely	Very likely	I don't know
☐	☐	☐	☐	☐

Part 4. Questions about how you feel about school

Instructions: For the following questions, there are no right or wrong answers. For each question, choose what you think is the best answer and write the letter for that answer in the box provided.

27. How much do you like school?

Options:

A) I dislike school a lot
B) I dislike school
C) I like school
D) I like school a lot

Answer: ☐

28. How helpful to you is what you are learning in school?

Options:

A) I think it is very unhelpful
B) I think it is unhelpful
C) I think it is helpful
D) I think it is very helpful

Answer: ☐

29. How much do you like science as a subject?

Options:

A) I dislike science a lot
B) I dislike science
C) I like science
D) I like science a lot

Answer: ☐

30. How helpful to you is what you are learning in science?

Options:

A) I think it is very unhelpful
B) I think it is unhelpful
C) I think it is helpful
D) I think it is very helpful

Answer: ☐

31. How much did you like what you learned as part of the lessons with "The Health Choices Book"?

Options:

A) I did not like the lessons at all
B) I did not like the lessons
C) I liked the lessons
D) liked the lessons very much

Answer: ☐

32. How helpful to you is what you learned as part of the lessons with "The Health Choices Book"?

Options:

A) I think it is very unhelpful
B) I think it is unhelpful
C) I think it is helpful
D) I think it is very helpful

Answer: ☐

33. How easy or difficult to understand did you find the lessons with "The Health Choices Book"?

Options:

A) Very difficult to understand
B) Difficult to understand
C) Easy to understand
D) Very easy to understand

Answer:

34. Do you trust what you learned as part of the lessons with "The Health Choices Book"?

Options:

A) I have very little trust in it
B) I have little trust in it
C) I trust it
D) I trust it very much

Answer:

Appendix 3: Data dictionary for the children's English Claim questionnaire

Description of database:	Data from the evaluation of the evaluation of the Informed Health Choices (IHC) primary school intervention done in Uganda in 2016. The study evaluated the effect of using the IHC primary school materials on children's understanding and ability to apply key concepts to assess claims about the effects (benefits and harms) of treatments. The intervention included two days of training for the teachers. We used multiple-choice questions from the Claim Evaluation Tools database in a test to measure the outcome. Each test was entered twice by two independent individuals (double data entry) and the whole dataset was checked for consistency up to generation of this final clean version. Some children had parents who were enrolled in a separate but related trial evaluating the effect of a podcast on parents' understanding and ability to apply key concepts to assess claims about treatment effects.
Number of variables:	58

No	Variable name	Type of variable	Description of variable data: Responses to:	Number of Characters	Values and value labels
1	recordid	Numerical	Number associated with a record		
2	Qn0part1	Numerical	Child's id in class	3	
3	Qn0part2	Alphanumeric	School id	4	
4	Qn0part3	Alphanumeric	Teacher's id	7	
5	Qn0part4	Alphanumeric	Child's study number	12	
6	Qn0part5		Child's school study group	1	1= Group 1, 2= Group 2
7	Qn0part6	Numerical	Parent enrolled in podcast trial?	1	1 = Enrolled, 2 = Not Enrolled
8	Qn0part7	Alphanumeric	If yes to Qn0part6 above, parent's study number	12	
9	Qn0part8	Alphanumeric	If yes to Qn0part6 above, parent's study group	1	1 = Group 1, 2 = Group 2
10	Qn0part9	Numerical	School ownership	1	1= Government 2= Private
11	Qn0part10	Numerical	School location	1	1=Rural 2=Semi urban 3=Urban
Part 2: Questions about the child					
12	Qn1part2		Child's age in completed years	2	0 = missing
13	Qn1part3		Child's gender	1	1 = Girl, 2 = Boy
14	Qn2part1	String	Who went to the doctor's office every day?	1	A = People who didn't, B = People who got..., C = Everyone, D = Sick people..., 0 = Missing
15	Qn2part2	String	How did the doctor decide who should get tea?	1	A = By tossing..., B = By asking people, C = The doctor gave, D = The doctor asked, 0 = Missing

No	Variable name	Type of variable	Description of variable data: Responses to:	Number of Characters	Values and value labels
16	Qn3part1	String	What was the treatment?	1	A = Tea, B = Sleep, C = The study, D = The doctor, 0 = Missing
17	Qn3part2	String	What was the result of the study?	1	A = Drinking tea can, B = Doctors toss coin, C = People should go, D = Not drinking tea, 0 = Missing
18	Qn4	String	Annette sees a soap advert. Thinks new soap is better. Is Annette right?	1	A = No, the soap..., B = Yes, soap is new, C = Yes, company is well-known, 0 = Missing
19	Qn5	String	Regina has an illness… How sure can she be that the medicine has no harms?	1	A = It is not..., B = Not very sure..., C = Very sure, since, 0 = Missing
20	Qn6	String	John has skin rash. He chooses a cream from a known company Is John right?	1	A = No, just because..., B = It is not possible..., C = No, the cream..., D = Yes, the company..., 0 = Missing
20	Qn7	String	Two companies make medicines. How can you know which medicine is better?	1	A = It isn't possible, B = I would rely..., C = I cannot trust, 0 = Missing
22	Qn8	String	Dr. Kato and Dr. Semakula disagree about which medicine. Is Dr. Kato right?	1	A = Yes, basis is experience, B = Yes, basis is studies, C = No, basis is not studies, 0 = Missing
23	Qn9	String	Habibah has pain in her ear. She asks Hassan. Do you agree with Hassan?	1	A = Yes, because this..., B = No, Hassan's experience, C = Yes, Hassan rinsed..., 0 = Missing
24	Qn10	String	Sarah has an illness. Three friends advise. Which is the best advice?	1	A = Only if many tried it, B = Only if it will not harm her, C = If Sarah has money, 0 = Missing
25	Qn11	String	Dr. Acheng is an expert. How sure can we be that Dr. Acheng is right?	1	A = It is not possible..., B = Not very sure..., C = Dr. Acheng is…, D = Dr. Acheng wouldn't be, 0 = Missing
26	Qn12	String	Edith has stomach pain. How sure can we be that juice is a good treatment?	1	A = Even though people..., B = Very sure..., C = Edith should ask..., 0 = Missing

No	Variable name	Type of variable	Description of variable data: Responses to:	Number of Characters	Values and value labels
27	Qn13	String	At David's school some children have poor parents. Is David correct?	1	A = It depends on..., B = Yes, students with..., C = Yes, the juice..., D = There could be other, 0 = Missing
28	Qn14	String	In a research study done by John… Mildred says we cannot be sure. Why?	1	A = Because the study..., B = Because John..., C = Because four people..., 0 = Missing
29	Qn15	String	A new fruit drink. Why can't we be sure about the results of Fred's study?	1	A = Because all people taking..., B = Because people knew if..., C = Because Fred should..., 0 = Missing
30	Qn16	String	Harriet is worried. How sure can she be that the old medicine is better…?	1	A = Harriet needs to..., B = She heard about..., C = Unless she finds..., D = This is a new study..., 0 = Missing
31	Qn17	String	Doctors studied people… Can we be sure that the new medicine is good…?	1	A = No taking the new..., B = Yes, people were asked..., C = Yes, the study was done..., 0 = Missing
32	Qn18	String	In a research study… How sure can you be the new treatment is better?	1	A = People taking the new..., B = All people taking part..., C = Older people didn't like..., 0 = Missing
33	Qn19	String	Judith wants smoother skin. Based on this link… Is Judith correct?	1	A = It depends on how..., B = There might be..., C = Yes, because the..., D = No, Judith should..., 0 = Missing
34	Qn20	String	Wasswa has done a research study giving. Is Dr. Wasswa right?	1	A = No. The people…, B = Yes, some of them, C = No, since not all, 0 = missing
35	Qn21part1	Numerical	Peter says that if a treatment works for one person	1	1 = I agree, 2 = I disagree, 0 = Missing
36	Qn21part2	Numerical	Alice says that if some people try the treatment and feel	1	1 = I agree, 2 = I disagree, 0 = Missing
37	Qn21part3	Numerical	Habibah says that just because many people are using the…	1	1 = I agree, 2 = I disagree, 0 = Missing
38	Qn21part4	Numerical	Julie says that companies sometimes say that the treatment…	1	1 = I agree, 2 = I disagree, 0 = Missing

No	Variable name	Type of variable	Description of variable data: Responses to:	Number of Characters	Values and value labels
39	Qn22part1	Numerical	The doctor should use chance to decide which people…	1	1 = More sure, 2 = Less sure, 0 = Missing
40	Qn22part2	Numerical	People should not know which medicine they get until…	1	1 = More sure, 2 = Less sure, 0 = Missing
41	Qn22part3	Numerical	The doctor should include only a small number of people…	1	1 = More sure, 2 = Less sure, 0 = Missing
42	Qn23part1	Numerical	Julie says that, if a treatment has been compared in a study…	1	1 = I agree, 2 = I disagree, 0 = Missing
43	Qn23part2	Numerical	Margaret says that the results of a study should be used to…	1	1 = I agree, 2 = I disagree, 0 = Missing
Part 3: Questions about your views					
44	Qn24part1	Numerical	Find out what the claim was based on, for example…	1	1 = Very unlikely, 2 = Unlikely, 3 = Likely, 4 = Very likely, 5 = I don't know, 0 = Missing
45	Qn24part2	Numerical	Find out if the claim was based on a research study comparing…	1	1 = Very unlikely, 2 = Unlikely, 3 = Likely, 4 = Very likely, 5 = I don't know, 0 = Missing
46	Qn25part1	Numerical	Assessing whether a claim about a treatment is based on a…	1	1 = Very difficult, 2 = Difficult, 3 = Easy, 4 = Very easy, 5 = I don't know, 0 = Missing
47	Qn25part2	Numerical	Assessing where I can find information about treatments…	1	1 = Very difficult, 2 = Difficult, 3 = Easy, 4 = Very easy, 5 = I don't know, 0 = Missing
48	Qn25part3	Numerical	Assessing how sure I can be about the results of a research…	1	1 = Very difficult, 2 = Difficult, 3 = Easy, 4 = Very easy, 5 = I don't know, 0 = Missing
49	Qn25part4	Numerical	Assessing if the results of a research study comparing…	1	1 = Very difficult, 2 = Difficult, 3 = Easy, 4 = Very easy, 5 = I don't know, 0 = Missing

No	Variable name	Type of variable	Description of variable data: Responses to:	Number of Characters	Values and value labels
50	Qn26	Numerical	Think about an illness you might get. How likely are you to say yes?	1	1 = Very unlikely, 2 = Unlikely, 3 = Likely, 4 = Very likely, 5 = I don't know, 0 = Missing
Questions about school					
51	Qn27	String	How much do you like school?	1	A = I dislike school a lot, B = I dislike school, C = I like school, D = I like school a lot
52	Qn28	String	How helpful is what you are learning in school?	1	A = I think it is very unhelpful, B = I think it is unhelpful, C = I think it is helpful, D = I think it is very helpful
53	Qn29	String	How much do like science as a subject?	1	A = I dislike science a lot, B = I dislike science, C = I like science, D = I like science a lot
54	Qn30	String	How helpful to you is what you are learning in science?	1	A = I think it is very unhelpful, B = I think it is unhelpful, C = I think it is helpful, D = I think it is very helpful
55	Qn31	String	How much did you like what you learned as part of the IHC lessons?	1	A = I did not like the lessons at all, B = I did not like the lessons, C = I liked the lessons, D = I liked the lessons very much
56	Qn32	String	How helpful to you is what you learned as part of the IHC lessons?	1	A = I think it is very unhelpful, B = I think it is unhelpful, C = I think it is helpful, D = I think it is very helpful
57	Qn33	String	How easy or difficult to understand did you find the IHC lessons?	1	A = Very difficult to understand, B = Difficult to understand, C = Easy to understand, D = Very easy to understand
58	Qn34	String	Do you trust what you learned as part of the lessons IHC lessons?	1	A = I have very little trust in it, B = I have little trust in it, C = I trust it, D = I trust it very much

End of questionnaire

Appendix 4: Data dictionary for the teachers' Claim questionnaire

Description of database:	Data from the evaluation of the Informed Health Choices (IHC) primary school intervention done in Uganda in 2016. The study evaluated the effect of using the IHC primary school materials on children's understanding and ability to apply key concepts to assess claims about the effects (benefits and harms) of treatments. We used the Claim Evaluation Tools database to measure the outcome. The dataset described here is from the test completed by teachers of the schools which participated in the IHC trials. Each test was entered twice by two independent individuals (double data entry) and the whole dataset was checked for consistency.
Number of variables:	81

No	Variable name	Type of variable	Description of variable data: Responses to:	Number of Characters	Values and value labels
1	Record id	Numerical	Unique id (primary key)	3	-
2	Qn0part1	Alphanumeric	School id	4	-
3	Qn0part2	Alphanumeric	Teacher's study number (primary key)	7	-
4	Qn0part3	Numerical	School study group	1	1 = Group 1, 2 = Group 2
5	Qn0part4	Numerical	Average class size	3	
6	Qn0part5	Numerical	Main subject taught by the teacher	1	1= Science, 2= Mathematics, 3 = Social Studies, 4 = English, 5 = Religious Education, 6 = Music, 7 = Computer, 8 = Others, 0 = Missing
7	Qn0part6	Numerical	Number of additional subjects	2	-
8	Qn0part7	Numerical	Number of classes or streams taught	2	-
9	Qn0part8	Numerical	Number of lessons per week	2	-
10	Qn0part9	Numerical	Additional schools taught?	1	1 = Yes, 2 = No, 0 = Missing
11	Qn0part10	Numerical	Age	2	0 = Missing
12	Qn0part11	Numerical	School ownership		1= Government 2 = Private
13	Qn0part12	Numerical	School location		1 =Rural 2 = Semi-urban 3 = Urban
Part 2: Questions about the teacher					
14	Qn1part2	Numerical	Gender	1	1 = Female, 2 = Male, 0 = Missing
15	Qn1part3	Numerical	Highest level of education	1	1 = Grade 3(Certificate), 2 = Grade 5(Diploma), 3 = University degree, 0 = Missing
16	Qn1part4	Numerical	Years of teaching experience	2	
17	Qn1part5	Numerical	Years taught science	2	

No	Variable name	Type of variable	Description of variable data: Responses to:	Number of Characters	Values and value labels
18	Qn1part6	Numerical	Had any training in research?	1	1 = Yes, 2 = No, 0 = Missing
19	Qn1part7	Numerical	Ever been a participant in research before?	1	1 = Yes, 2 = No, 0 = Missing
Part 3. Questions about assessing claims					
20	Qn2part1	Categorical - string	Who went to the doctor's office every day?	1	A = People who didn't, B = People who got..., C = Everyone, D = Sick people..., 0 = Missing
21	Qn2part2	String	How did the doctor decide who should get tea?	1	A = By tossing..., B = By asking people, C = The doctor gave, D = The doctor asked, 0 = Missing
22	Qn3part1		What was the treatment?	1	A = Tea, B = Sleep, C = The study, D = The doctor, 0 = Missing
23	Qn3part2		What was the result of the study?	1	A = Drinking tea can, B = Doctors toss coin, C = People should go, D = Not drinking tea, 0 = Missing
24	Qn4		Annette sees a soap advert. Thinks new soap is better. Is Annette right?	1	A = No, the soap..., B = Yes, soap is new, C = Yes, company is well-known, 0 = Missing
25	Qn5		Regina has an illness… How sure can she be that the medicine has no harms?	1	A = It is not..., B = Not very sure..., C = Very sure, since, 0 = Missing
26	Qn6		John has skin rash. He chooses a cream from a known company Is John right?	1	A = No, just because..., B = It is not possible..., C = No, the cream..., D = Yes, the company..., 0 = Missing
27	Qn7		Two companies make medicines. How can you know which medicine is better?	1	A = It isn't possible, B = I would rely..., C = I cannot trust, 0 = Missing
28	Qn8		Dr. Kato and Dr. Semakula disagree about which medicine. Is Dr. Kato right?	1	A = Yes, basis is experience, B = Yes, basis is studies, C = No, basis is not studies, 0 = Missing

No	Variable name	Type of variable	Description of variable data: Responses to:	Number of Characters	Values and value labels
29	Qn9		Habibah has pain in her ear. She asks Hassan. Do you agree with Hassan?	1	A = Yes, because this..., B = No, Hassan's experience, C = Yes, Hassan rinsed..., 0 = Missing
30	Qn10		Sarah has an illness. Three friends advise. Which is the best advice?	1	A = Only if many tried it, B = Only if it will not harm her, C = If Sarah has money, 0 = Missing
31	Qn11		Dr. Acheng is an expert. How sure can we be that Dr. Acheng is right?	1	A = It is not possible..., B = Not very sure..., C = Dr. Acheng is…, D = Dr. Acheng wouldn't be, 0 = Missing
32	Qn12		Edith has stomach pain. How sure can we be that juice is a good treatment?	1	A = Even though people..., B = Very sure..., C = Edith should ask..., 0 = Missing
33	Qn13		At David's school some children have poor parents. Is David correct?	1	A = It depends on..., B = Yes, students with..., C = Yes, the juice..., D = There could be other, 0 = Missing
34	Qn14		In a research study done by John… Mildred says we cannot be sure. Why?	1	A = Because the study..., B = Because John..., C = Because four people..., 0 = Missing
35	Qn15		A new fruit drink. Why can't we be sure about the results of Fred's study?	1	A = Because all people taking..., B = Because people knew if..., C = Because Fred should..., 0 = Missing
36	Qn16		Harriet is worried. How sure can she be that the old medicine is better…?	1	A = Harriet needs to..., B = She heard about..., C = Unless she finds..., D = This is a new study..., 0 = Missing
37	Qn17		Doctors studied people… Can we be sure that the new medicine is good…?	1	A = No taking the new..., B = Yes, people were asked..., C = Yes, the study was done..., 0 = Missing

No	Variable name	Type of variable	Description of variable data: Responses to:	Number of Characters	Values and value labels
38	Qn18		In a research study… How sure can you be the new treatment is better?	1	A = People taking the new..., B = All people taking part..., C = Older people didn't like..., 0 = Missing
39	Qn19		Judith wants smoother skin. Based on this link… Is Judith correct?	1	A = It depends on how..., B = There might be..., C = Yes, because the..., D = No, Judith should..., 0 = Missing
40	Qn20		Wasswa has done a research study giving. Is Dr. Wasswa right?	1	A = No. The people…, B = Yes, some of them, C = No, since not all, 0 = missing
41	Qn21part1	Numerical	Peter says that if a treatment works for one person	1	1 = I agree, 2 = I disagree, 0 = Missing
42	Qn21part2	Numerical	Alice says that if some people try the treatment and feel	1	1 = I agree, 2 = I disagree, 0 = Missing
43	Qn21part3	Numerical	Habibah says that just because many people are using the…	1	1 = I agree, 2 = I disagree, 0 = Missing
44	Qn21part4	Numerical	Julie says that companies sometimes say that the treatment…	1	1 = I agree, 2 = I disagree, 0 = Missing
45	Qn22part1	Numerical	The doctor should use chance to decide which people…	1	1 = More sure, 2 = Less sure, 0 = Missing
46	Qn22part2	Numerical	People should not know which medicine they get until…	1	1 = More sure, 2 = Less sure, 0 = Missing
47	Qn22part3	Numerical	The doctor should include only a small number of people…	1	1 = More sure, 2 = Less sure, 0 = Missing
48	Qn23part1	Numerical	Julie says that, if a treatment has been compared in a study…	1	1 = I agree, 2 = I disagree, 0 = Missing
49	Qn23part2	Numerical	Margaret says that the results of a study should be used to…	1	1 = I agree, 2 = I disagree, 0 = Missing
Part 4: Questions about the teachers' self-assessed competencies					
50	Qn24part1	Numerical	Find out what the claim was based on, for example…	1	1 = Very unlikely, 2 = Unlikely, 3 = Likely, 4 = Very likely, 5 = I don't know, 0 = Missing

No	Variable name	Type of variable	Description of variable data: Responses to:	Number of Characters	Values and value labels
51	Qn24part2	Numerical	Find out if the claim was based on a research study comparing…	1	1 = Very unlikely, 2 = Unlikely, 3 = Likely, 4 = Very likely, 5 = I don't know, 0 = Missing
52	Qn25part1	Numerical	Assessing whether a claim about a treatment is based on a…	1	1 = Very difficult, 2 = Difficult, 3 = Easy, 4 = Very easy, 5 = I don't know, 0 = Missing
53	Qn25part2	Numerical	Assessing where I can find information about treatments…	1	1 = Very difficult, 2 = Difficult, 3 = Easy, 4 = Very easy, 5 = I don't know, 0 = Missing
54	Qn25part3	Numerical	Assessing how sure I can be about the results of a research…	1	1 = Very difficult, 2 = Difficult, 3 = Easy, 4 = Very easy, 5 = I don't know, 0 = Missing
55	Qn25part4	Numerical	Assessing if the results of a research study comparing…	1	1 = Very difficult, 2 = Difficult, 3 = Easy, 4 = Very easy, 5 = I don't know, 0 = Missing
56	Qn26	Numerical	Think about an illness you might get. How likely are you to say yes?	1	1 = Very unlikely, 2 = Unlikely, 3 = Likely, 4 = Very likely, 5 = I don't know, 0 = Missing
Part 5. Additional questions completed by teachers of the schools in the intervention arm of the trial)					
57	Qn27	string	How important do you think fair comparisons of educational intervention…	1	A = I think they are very unimportant, B = I think they are unimportant, C = I think they are important, D = I think they are very important, 0 = Missing
58	Qn28	string	How much did you like teaching the lessons?	1	A = I did not like the lessons at all, B = I did not like the lessons, C = I liked the lessons, D = I liked the lessons very much, 0 = Missing

No	Variable name	Type of variable	Description of variable data: Responses to:	Number of Characters	Values and value labels
59	Qn29	string	How important is what the children learned from the lessons?	1	A = I think it is unimportant, B = I think it is of little importance, C = I think it is important, D = I think it is very important, 0 = Missing
60	Qn30	string	How easy or difficult was it for the children in your class to understand...	1	A = Very difficult to understand, B = Difficult to understand, C = Easy to understand, D = Very easy to understand, 0 = Missing
61	Qn31	string	How much do you think the children learned from the lessons?	1	A = I think they learned very little, B = I think they learned little, C = I think they learned much, D = I think they learned very much, 0 = Missing
62	Qn32	string	How much do you think you learned from the lessons?	1	A = I think I learned very little, B = I think I learned little, C = I think I learned much, D = I think I learned very much, 0 = Missing
63	qn33part1	Numerical	The training that I received on how to teach the lessons was ...	1	1 = Strongly disagree, 2 = Disagree, 3 = Neither agree nor disagree, 4 = Agree, 5 = Strongly agree, 0 = Missing
64	qn33part2	Numerical	The training that I received on how to teach the lessons was	1	1 = Strongly disagree, 2 = Disagree, 3 = Neither agree nor disagree, 4 = Agree, 5 = Strongly agree, 0 = Missing
65	qn33part3	Numerical	I understood the content of the lessons well	1	1 = Strongly disagree, 2 = Disagree, 3 = Neither agree nor disagree, 4 = Agree, 5 = Strongly agree, 0 = Missing

No	Variable name	Type of variable	Description of variable data: Responses to:	Number of Characters	Values and value labels
66	qn33part4	Numerical	I am confident about my ability to teach the lessons to the children	1	1 = Strongly disagree, 2 = Disagree, 3 = Neither agree nor disagree, 4 = Agree, 5 = Strongly agree, 0 = Missing
67	qn33part5	Numerical	The teacher's guide was useful in conducting the lessons	1	1 = Strongly disagree, 2 = Disagree, 3 = Neither agree nor disagree, 4 = Agree, 5 = Strongly agree, 0 = Missing
68	qn33part6	Numerical	The instructions for how to teach the lessons fit my teaching style ...	1	1 = Strongly disagree, 2 = Disagree, 3 = Neither agree nor disagree, 4 = Agree, 5 = Strongly agree, 0 = Missing
69	qn33part7	Numerical	The lesson activities helped children understand the content better	1	1 = Strongly disagree, 2 = Disagree, 3 = Neither agree nor disagree, 4 = Agree, 5 = Strongly agree, 0 = Missing
70	qn33part8	Numerical	I had to adapt the instructions to fit my teaching style	1	1 = Strongly disagree, 2 = Disagree, 3 = Neither agree nor disagree, 4 = Agree, 5 = Strongly agree, 0 = Missing
71	qn33part9	Numerical	I liked teaching something new	1	1 = Strongly disagree, 2 = Disagree, 3 = Neither agree nor disagree, 4 = Agree, 5 = Strongly agree, 0 = Missing
72	qn33part10	Numerical	I think it is important to teach children to think critically	1	1 = Strongly disagree, 2 = Disagree, 3 = Neither agree nor disagree, 4 = Agree, 5 = Strongly agree, 0 = Missing

No	Variable name	Type of variable	Description of variable data: Responses to:	Number of Characters	Values and value labels
73	qn33part11	Numerical	I do not like being challenged by children in my class	1	1 = Strongly disagree, 2 = Disagree, 3 = Neither agree nor disagree, 4 = Agree, 5 = Strongly agree, 0 = Missing
74	qn33part12	Numerical	I like the content of the lessons	1	1 = Strongly disagree, 2 = Disagree, 3 = Neither agree nor disagree, 4 = Agree, 5 = Strongly agree, 0 = Missing
75	qn33part13	Numerical	I liked the way the teaching materials and lessons were organised	1	1 = Strongly disagree, 2 = Disagree, 3 = Neither agree nor disagree, 4 = Agree, 5 = Strongly agree, 0 = Missing
76	qn33part14	Numerical	I found teaching the lessons stressful	1	1 = Strongly disagree, 2 = Disagree, 3 = Neither agree nor disagree, 4 = Agree, 5 = Strongly agree, 0 = Missing
77	qn33part15	Numerical	I felt motivated to teach the lessons	1	1 = Strongly disagree, 2 = Disagree, 3 = Neither agree nor disagree, 4 = Agree, 5 = Strongly agree, 0 = Missing
78	qn33part16	Numerical	I like to be asked questions by children in my class and for there ...	1	1 = Strongly disagree, 2 = Disagree, 3 = Neither agree nor disagree, 4 = Agree, 5 = Strongly agree, 0 = Missing
79	qn33part17	Numerical	There was not enough time to teach the lessons	1	1 = Strongly disagree, 2 = Disagree, 3 = Neither agree nor disagree, 4 = Agree, 5 = Strongly agree, 0 = Missing

No	Variable name	Type of variable	Description of variable data: Responses to:	Number of Characters	Values and value labels
80	qn33part18	Numerical	The class timetable was flexible enough to include the lessons	1	1 = Strongly disagree, 2 = Disagree, 3 = Neither agree nor disagree, 4 = Agree, 5 = Strongly agree, 0 = Missing
81	qn33part19	Numerical	The head teacher thinks it is important to teach the lessons	1	1 = Strongly disagree, 2 = Disagree, 3 = Neither agree nor disagree, 4 = Agree, 5 = Strongly agree, 0 = Missing

End of questionnaire

Appendix 5: Data dictionary for the children's Luganda Claim questionnaire

Description of database:	Data from the evaluation of the evaluation of the Informed Health Choices (IHC) primary school intervention done in Uganda in 2016. The study evaluated the effect of the IHC primary school intervention on children's understanding and ability to apply key concepts to assess claims about the effects (benefits and harms) of treatments. We used multiple-choice questions from the Claim Evaluation Tools database in a test to measure the outcome. Each test was entered twice by two independent individuals (double data entry) and the whole dataset was checked for consistency up to generation of this final clean version. Some children had parents who were enrolled in a separate but related trial evaluating the effect of a podcast on parents' understanding and ability to apply key concepts to assess claims about treatment effects.
Number of variables:	58

No	Variable name	Type of variable	Description of variable data: Responses to:	Number of Characters	Values and value labels
1	recordid	Numerical	Number associated with a record		
2	Qn0part1	Numerical	Child's id in class	3	
3	Qn0part2	Alphanumeric	School id	4	
4	Qn0part3	Alphanumeric	Teacher's id	7	
5	Qn0part4	Alphanumeric	Child's study number	12	
6	Qn0part5		Child's school study group	1	1= Group 1, 2= Group 2
7	Qn0part6	Numerical	Parent enrolled in podcast trial?	1	1 = Enrolled, 2 = Not Enrolled
8	Qn0part7	Alphanumeric	If yes to Qn0part6 above, parent's study number	12	
9	Qn0part8	Alphanumeric	If yes to Qn0part6 above, parent's study group	1	1 = Group 1, 2 = Group 2
10	Qn0part9	Numerical	School ownership	1	1= Government 2= Private
11	Qn0part10	Numerical	School location	1	1=Rural 2=Semi urban 3=Urban
Part 2: Questions about the child					
12	Qn1part2		Child's age in completed years	2	0 = missing
13	Qn1part3		Child's gender	1	1 = Girl, 2 = Boy
14	Qn2part1	String	Who went to the doctor's office every day?	1	A = People who didn't, B = People who got..., C = Everyone, D = Sick people..., 0 = Missing
15	Qn2part2	String	How did the doctor decide who should get tea?	1	A = By tossing..., B = By asking people, C = The doctor gave, D = The doctor asked, 0 = Missing
16	Qn3part1	String	What was the treatment?	1	A = Tea, B = Sleep, C = The study, D = The doctor, 0 = Missing

No	Variable name	Type of variable	Description of variable data: Responses to:	Number of Characters	Values and value labels
17	Qn3part2	String	What was the result of the study?	1	A = Drinking tea can, B = Doctors toss coin, C = People should go, D = Not drinking tea, 0 = Missing
18	Qn4	String	Annette sees a soap advert. Thinks new soap is better. Is Annette right?	1	A = No, the soap..., B = Yes, soap is new, C = Yes, company is well-known, 0 = Missing
19	Qn5	String	Regina has an illness… How sure can she be that the medicine has no harms?	1	A = It is not..., B = Not very sure..., C = Very sure, since, 0 = Missing
20	Qn6	String	John has skin rash. He chooses a cream from a known company Is John right?	1	A = No, just because..., B = It is not possible..., C = No, the cream..., D = Yes, the company..., 0 = Missing
20	Qn7	String	Two companies make medicines. How can you know which medicine is better?	1	A = It isn't possible, B = I would rely..., C = I cannot trust, 0 = Missing
22	Qn8	String	Dr. Kato and Dr. Semakula disagree about which medicine. Is Dr. Kato right?	1	A = Yes, basis is experience, B = Yes, basis is studies, C = No, basis is not studies, 0 = Missing
23	Qn9	String	Habibah has pain in her ear. She asks Hassan. Do you agree with Hassan?	1	A = Yes, because this..., B = No, Hassan's experience, C = Yes, Hassan rinsed..., 0 = Missing
24	Qn10	String	Sarah has an illness. Three friends advise. Which is the best advice?	1	A = Only if many tried it, B = Only if it will not harm her, C = If Sarah has money, 0 = Missing
25	Qn11	String	Dr. Acheng is an expert. How sure can we be that Dr. Acheng is right?	1	A = It is not possible..., B = Not very sure..., C = Dr. Acheng is…, D = Dr. Acheng wouldn't be, 0 = Missing
26	Qn12	String	Edith has stomach pain. How sure can we be that juice is a good treatment?	1	A = Even though people..., B = Very sure..., C = Edith should ask..., 0 = Missing
27	Qn13	String	At David's school some children have poor parents. Is David correct?	1	A = It depends on..., B = Yes, students with..., C = Yes, the juice..., D = There could be other, 0 = Missing

No	Variable name	Type of variable	Description of variable data: Responses to:	Number of Characters	Values and value labels
28	Qn14	String	In a research study done by John… Mildred says we cannot be sure. Why?	1	A = Because the study..., B = Because John..., C = Because four people..., 0 = Missing
29	Qn15	String	A new fruit drink. Why can't we be sure about the results of Fred's study?	1	A = Because all people taking..., B = Because people knew if..., C = Because Fred should..., 0 = Missing
30	Qn16	String	Harriet is worried. How sure can she be that the old medicine is better…?	1	A = Harriet needs to..., B = She heard about..., C = Unless she finds..., D = This is a new study..., 0 = Missing
31	Qn17	String	Doctors studied people… Can we be sure that the new medicine is good…?	1	A = No taking the new..., B = Yes, people were asked..., C = Yes, the study was done..., 0 = Missing
32	Qn18	String	In a research study… How sure can you be the new treatment is better?	1	A = People taking the new..., B = All people taking part..., C = Older people didn't like..., 0 = Missing
33	Qn19	String	Judith wants smoother skin. Based on this link… Is Judith correct?	1	A = It depends on how..., B = There might be..., C = Yes, because the..., D = No, Judith should..., 0 = Missing
34	Qn20	String	Wasswa has done a research study giving. Is Dr. Wasswa right?	1	A = No. The people…, B = Yes, some of them, C = No, since not all, 0 = missing
35	Qn21part1	Numerical	Peter says that if a treatment works for one person	1	1 = I agree, 2 = I disagree, 0 = Missing
36	Qn21part2	Numerical	Alice says that if some people try the treatment and feel	1	1 = I agree, 2 = I disagree, 0 = Missing
37	Qn21part3	Numerical	Habibah says that just because many people are using the…	1	1 = I agree, 2 = I disagree, 0 = Missing
38	Qn21part4	Numerical	Julie says that companies sometimes say that the treatment…	1	1 = I agree, 2 = I disagree, 0 = Missing
39	Qn22part1	Numerical	The doctor should use chance to decide which people…	1	1 = More sure, 2 = Less sure, 0 = Missing
40	Qn22part2	Numerical	People should not know which medicine they get until…	1	1 = More sure, 2 = Less sure, 0 = Missing

No	Variable name	Type of variable	Description of variable data: Responses to:	Number of Characters	Values and value labels
41	Qn22part3	Numerical	The doctor should include only a small number of people…	1	1 = More sure, 2 = Less sure, 0 = Missing
42	Qn23part1	Numerical	Julie says that, if a treatment has been compared in a study…	1	1 = I agree, 2 = I disagree, 0 = Missing
43	Qn23part2	Numerical	Margaret says that the results of a study should be used to…	1	1 = I agree, 2 = I disagree, 0 = Missing
Part 3: Questions about your views					
44	Qn24part1	Numerical	Find out what the claim was based on, for example…	1	1 = Very unlikely, 2 = Unlikely, 3 = Likely, 4 = Very likely, 5 = I don't know, 0 = Missing
45	Qn24part2	Numerical	Find out if the claim was based on a research study comparing…	1	1 = Very unlikely, 2 = Unlikely, 3 = Likely, 4 = Very likely, 5 = I don't know, 0 = Missing
46	Qn25part1	Numerical	Assessing whether a claim about a treatment is based on a…	1	1 = Very difficult, 2 = Difficult, 3 = Easy, 4 = Very easy, 5 = I don't know, 0 = Missing
47	Qn25part2	Numerical	Assessing where I can find information about treatments…	1	1 = Very difficult, 2 = Difficult, 3 = Easy, 4 = Very easy, 5 = I don't know, 0 = Missing
48	Qn25part3	Numerical	Assessing how sure I can be about the results of a research…	1	1 = Very difficult, 2 = Difficult, 3 = Easy, 4 = Very easy, 5 = I don't know, 0 = Missing
49	Qn25part4	Numerical	Assessing if the results of a research study comparing…	1	1 = Very difficult, 2 = Difficult, 3 = Easy, 4 = Very easy, 5 = I don't know, 0 = Missing
50	Qn26	Numerical	Think about an illness you might get. How likely are you to say yes?	1	1 = Very unlikely, 2 = Unlikely, 3 = Likely, 4 = Very likely, 5 = I don't know, 0 = Missing

No	Variable name	Type of variable	Description of variable data: Responses to:	Number of Characters	Values and value labels
Questions about school					
51	Qn27	String	How much do you like school?	1	A = I dislike school a lot, B = I dislike school, C = I like school, D = I like school a lot
52	Qn28	String	How helpful is what you are learning in school?	1	A = I think it is very unhelpful, B = I think it is unhelpful, C = I think it is helpful, D = I think it is very helpful
53	Qn29	String	How much do like science as a subject?	1	A = I dislike science a lot, B = I dislike science, C = I like science, D = I like science a lot
54	Qn30	String	How helpful to you is what you are learning in science?	1	A = I think it is very unhelpful, B = I think it is unhelpful, C = I think it is helpful, D = I think it is very helpful
55	Qn31	String	How much did you like what you learned as part of the IHC lessons?	1	A = I did not like the lessons at all, B = I did not like the lessons, C = I liked the lessons, D = I liked the lessons very much
56	Qn32	String	How helpful to you is what you learned as part of the IHC lessons?	1	A = I think it is very unhelpful, B = I think it is unhelpful, C = I think it is helpful, D = I think it is very helpful
57	Qn33	String	How easy or difficult to understand did you find the IHC lessons?	1	A = Very difficult to understand, B = Difficult to understand, C = Easy to understand, D = Very easy to understand
58	Qn34	String	Do you trust what you learned as part of the lessons IHC lessons?	1	A = I have very little trust in it, B = I have little trust in it, C = I trust it, D = I trust it very much

End of questionnaire

Appendix 6: Attrition, differences in test scores across strata of schools, and missing values

Proportion of children who completed the test

Control schools				
	Government	**Private**	**Total**	**N (%) of schools**
Rural	67.4%	61.8%	65.8%	8 (13%)
Semi-urban	70.2%	70.2%	70.2%	15 (25%)
Urban	73.1%	70.5%	71.6%	37 (62%)
Total	**71.2%**	**69.8%**	**70.6%**	

Intervention schools				
	Government	**Private**	**Total**	**N (%) of schools**
Rural	74.8%	77.3%	76.3%	6 (10%)
Semi-urban	86.3%	93.1%	88.7%	14 (23%)
Urban	93.2%	92.5%	92.8%	40 (67%)
Total	**89.8%**	**91.1%**	**90.4%**	

Test scores

	Treatment effect	School ownership effect	School location effect
Mean score			
Without weighting	20.0% (95% CI 17.3 to 22.7) P < 0.00001	5.86% (95% CI 3.11% to 8.61%) P = 0.00003	1.24% (95% CI -0.63 to 3.12) P = 0.194
Weighted	20.0% (95% CI 17.3 to 22.7)	5.83% (95% CI: 3.08% to 8.57%) P = 0.00003	1.23% (95% CI -0.64 to 3.10) P = 0.199
Passing score (≥ 13 out of 24 correct answers)			
Without weighting	OR 9.34 (95% CI 6.62 to 13.18) (P < 0.00001)	OR1.98 (95% CI 1.40 to 2.81) P = 0.00011	OR 1.28 (95% CI 1.00 to 1.63) P = 0.04623
Weighted	OR 9.48 (95% CI 6.70 to 13.41) P < 0.00001	OR 1.97 (95% CI 1.39 to 2.80) P = 0.00015	OR 1.28 (95% CI 1.00 to 1.63) P = 0.04811
Mastery score (≥ 20 out of 24 correct answers)			
Without weighting	OR 35.33 (95% CI 20.58 to 60.67) P < 0.00001	OR 2.45 (95% CI 1.57 to 3.82) P = 0.00008	OR 1.07 (95% CI 0.76 to 1.51) P = 0.692
Weighted	OR 39.07 (95% CI 22.92 to 66.61) P < 0.00001	OR 2.55 (95% CI 1.61 to 4.03) P = 0.00006	OR 1.07 (95% CI 0.76 to 1.51) P = 0.706

OR: odds ratio

Number of missing values in each question for students

Number of unanswered questions
(Total number of completed tests: 6383 for intervention and 4430 for control schools)

	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11	Q12	Q13	Q14
Intervention	41	41	24	46	56	56	54	44	99	90	48
Control	34	22	25	24	21	23	30	19	36	27	30

	Q15	Q16	Q17	Q18	Q19	Q20	Q21 part1	Q21 part2	Q21 part3	Q21 part4	Q22 part1
Intervention	54	71	85	56	49	66	105	149	137	146	107
Control	30	33	52	26	20	38	144	169	224	203	126

	Q22 part2	Q22 part3	Q23 part1	Q23 part2	Q24 part1	Q24 part2	Q25 part1	Q25 part2	Q25 part3	Q25 part4
Intervention	114	127	81	84	189	211	171	202	212	194
Control	160	170	133	125	211	216	190	208	221	194

III

BMJ Open

Informed health choices intervention to teach primary school children in low-income countries to assess claims about treatment effects: process evaluation

Allen Nsangi,[1,2] Daniel Semakula,[1,2] Claire Glenton,[3] Simon Lewin,[3,4] Andrew D Oxman,[2,3] Matt Oxman,[3] Sarah Rosenbaum,[3] Astrid Dahlgren,[5] Laetitia Nyirazinyoye,[6] Margaret Kaseje,[7] Christopher James Rose,[3] Atle Fretheim,[2,3] Nelson K Sewankambo[1]

To cite: Nsangi A, Semakula D, Glenton C, *et al*. Informed health choices intervention to teach primary school children in low-income countries to assess claims about treatment effects: process evaluation. *BMJ Open* 2019;**9**:e030787. doi:10.1136/bmjopen-2019-030787

► Prepublication history and additional material for this paper are available online. To view these files, please visit the journal online (http://dx.doi.org/10.1136/bmjopen-2019-030787).

Received 12 April 2019
Revised 02 August 2019
Accepted 02 August 2019

For numbered affiliations see end of article.

Correspondence to
Dr Andrew D Oxman;
oxman@online.no

ABSTRACT

Background We developed the informed health choices (IHC) primary school resources to teach children how to assess the trustworthiness of claims about the effects of treatments. We evaluated these resources in a randomised trial in Uganda. This paper describes the process evaluation that we conducted alongside this trial.

Objectives To identify factors affecting the implementation, impact and scaling up of the intervention; and potential adverse and beneficial effects of the intervention.

Methods All 85 teachers in the 60 schools in the intervention arm of the trial completed a questionnaire after each lesson and at the end of the term. We conducted structured classroom observations at all 60 schools. For interviews and focus groups, we purposively selected six schools. We interviewed district education officers, teachers, head teachers, children and their parents. We used a framework analysis approach to analyse the data.

Results Most of the participants liked the IHC resources and felt that the content was important. This motivated the teachers and contributed to positive attitudes. Although some teachers started out lacking confidence, many found that the children's enthusiasm for the lessons made them more confident. Nearly everyone interviewed thought that the children learnt something important and many thought that it improved their decision-making. The main barrier to scaling up use of the IHC resources that participants identified was the need to incorporate the lessons into the national curriculum.

Conclusion The mostly positive findings reflect the trial results, which showed large effects on the children's and the teachers' critical appraisal skills. The main limitations of this evaluation are that the investigators were responsible for both developing and evaluating the intervention.

Strengths and limitations of this study

- ► Use of data collection triangulation.
- ► Having used a modified CERQual approach, we have high confidence in most of our findings.
- ► The study investigators were responsible for both developing and evaluating the intervention.

BACKGROUND

Adults and children are confronted with claims about the effects of treatments (any action intended to maintain or improve health) in their everyday interaction. Many of these claims are unsubstantiated, unreliable, inaccurate or biased.[1 2] Failure to use treatments supported by reliable evidence may result in unnecessary suffering and can waste scarce resources, especially in low-income countries. This could be avoided if people were able to assess the trustworthiness of treatment claims and make informed healthcare choices.[3 4] However, several studies have shown that people commonly lack the ability to understand the risks and benefits of treatments, and fail to apply key concepts that are essential for appraising claims about the effects of treatments.[5–8]

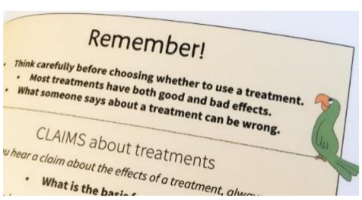
The aim of the informed health choices (IHC) project is to help address this problem by developing and evaluating learning resources to enable people to assess claims about treatment effects and make informed healthcare choices. We first developed a list of IHC Key Concepts that people need to understand in order to assess claims about the benefits and harms of treatments.[4] We determined which of these concepts could and should be taught to primary school children.[9] We then spent 3 years developing the IHC primary school resources[10] using a human centred design approach,[11] that included several cycles of idea generation, prototyping solutions, piloting in schools and making improvements grounded in teacher

The Health Choices Book
This textbook for primary school children (10 to 12 year olds) includes a comic story that introduces and explains 12 Key Concepts, instructions for classroom activities, exercises, a list of key messages, and a glossary.

Teachers' Guide
This guide includes lesson plans and other resources to help teachers using The Health Choices Book.

Exercise Book
This includes the exercises from The Health Choices Book.

Activity Cards
These cards are for use in Lesson 7 of The Health Choices Book. The activity is created to demonstrate how comparisons with few people can be misleading.



Checklist
This poster with the key messages from The Health Choices Book is a checklist for applying the 12 Key Concepts and a reminder of the most important messages in the book.

The 12 Key Concepts that are taught in the IHC primary school resources:

Recognising ***claims*** *about the effects of treatments that have an unreliable basis*

- Treatments may be harmful
- Personal experiences or anecdotes (stories) are an unreliable basis for assessing the effects of most treatments
- Widely used treatments or treatments that have been used for a long time are not necessarily beneficial or safe
- New, brand-named, or more expensive treatments may not be better than available alternatives
- Opinions of experts or authorities do not alone provide a reliable basis for deciding on the benefits and harms of treatments
- Conflicting interests may result in misleading claims about the effects of treatments

Understanding whether ***comparisons*** *of treatments are fair and reliable*

- Identifying effects of treatments depends on making comparisons
- Apart from the treatments being compared, the comparison groups need to be similar at the beginning of a comparison (i.e. 'like needs to be compared with like')
- If possible, people should not know which of the treatments being compared they are receiving
- Small studies in which few outcome events occur are usually not informative and the results may be misleading
- The results of single comparisons of treatments can be misleading

Making informed ***choices*** *about treatments*

- Decisions about treatments should not be based on considering only their benefits

Figure 1 The informed health choices primary school resources.

and student feedback and classroom observation. These resources facilitate the teaching of 12 of the IHC Key Concepts to grade 5 children (10–12 years) (figure 1).[9 10] We also developed a podcast to teach some of the same concepts to the children's parents.[12]

Teaching primary school children to assess claims about treatments can capitalise on children's curiosity and enthusiasm to learn.[13] Through children sharing what they are learning at school, it might also have an indirect effect on their family members' abilities to determine the

Nsangi A, *et al. BMJ Open* 2019;**9**:e030787. doi:10.1136/bmjopen-2019-030787

reliability of claims. Teaching children to assess claims about treatment effects can provide a foundation for a more scientifically literate and healthier society.

We evaluated the effects of IHC primary school intervention on children's ability to assess treatment claims in a cluster randomised trial in Uganda.[14] The trial showed that the intervention had a large effect on the children's ability to assess treatment claims. It also had a large effect on the teachers' ability to assess treatment claims. We measured these outcomes again after 1 year, and both children and teachers had retained what they learnt.[15] In a linked randomised trial, we evaluated the effects of the IHC podcast on the ability of parents of the primary school children to assess treatment claims.[16]

The objectives of this process evaluation were to:

1. Identify factors affecting the implementation, impact and scaling up of the intervention.
2. Identify potential adverse and beneficial effects of the intervention.

These objectives differ from the four objectives in the protocol for this study,[17] as described in online supplementary file 1. These changes were made in order to present the findings of both the qualitative and quantitative analyses more coherently and succinctly.

METHODS

This was a multimethod study using qualitative and quantitative data. Our main focus in this paper is on the qualitative analyses. Some of the quantitative results are reported elsewhere.[15]

Qualitative data collection and analysis

Frameworks

We used a framework thematic analysis approach to guide data collection and analysis.[17] We started out by developing two frameworks. The first addressed the factors that could affect the implementation, impact and scaling up of the school resources (table 1), and the second addressed potential adverse and beneficial effects of the resources (table 2).

The first framework (table 1) was developed iteratively by reviewing existing frameworks and studies of barriers and facilitators to implementing changes in schools,[18–22] and frameworks of barriers and facilitators to implement changes in health professional practice.[23] The framework shown in table 1 has been modified from the framework in the protocol, based on the results of the framework analysis.

Use of the IHC school resources might have had adverse or beneficial effects that were not measured in the trial, including effects on relationships between children and others, and on beliefs, attitudes or behaviours. We developed a framework of potential effects and impacts based on pilot and user testing of the resources; discussions with education researchers, policy-makers and teachers; potential beneficial effects identified by the National Curriculum Development Centre in Uganda, and reviewing the literature. The framework shown in table 2 has also been modified based on the results of the framework analysis.

Sampling

The intervention was implemented in 60 schools in Uganda.[14] For qualitative data, we sampled six of these schools. Within each school, we included all the head teachers, all the grade 5 teachers who participated in the IHC lessons, two children and two parents. In order to capture the opinions, views and experiences of a wide range of participants,[24] we purposively sampled the six schools based on geographical location (rural, semi-urban or urban) and ownership (public or private). We also sampled schools with variation in the extent to which teachers used the resources as intended. Within the schools, we sampled children with variation in performance on end-of-term examinations and based on how well they understood the IHC lessons. We sampled parents who had also participated in either the intervention or control group of the IHC podcast trial and who had varying levels of education. Finally, we aimed to include all of the five district education officers in the central region of Uganda, where the trial took place.

Data collection

We collected qualitative data using lesson evaluation forms, observation, individual interviews and focus group interviews (box 1).

The questions and prompts used in the interviews and focus group discussions were guided by the two frameworks described above. For all of the interviews and focus group discussions, one of the study investigators (AN or DS) carried out the interview or facilitated the discussion. A research assistant was responsible for observation and note taking during the interviews. All of the interviews and focus group discussions were conducted in English except for interviews with two head teachers, one teacher and all but one parent; which were done in Luganda. AN, DS and NKS are fluent in Luganda. We recorded and transcribed all of the interviews and focus group discussions. Observations during class lessons, interviews and focus groups were recorded using structured forms and entered into a spreadsheet.

We conducted a total of 44 individual interviews: 12 with children, 6 with head teachers, 10 with grade 5 teachers, 13 with parents and 3 with district education officers (2 individual interviews, 1 joint interview); and five focus group discussions (three with children and two with teachers). In addition, we observed at least two of the lessons taught in the six schools selected for the process evaluation. The amount of data we collected was guided by considerations of the variation in issues emerging from the data collection and the extent to which we are able to explain these variations; our time and resource constraints; and the need to avoid large volumes of data that cannot be easily managed or analysed as emphasised in literature.[24]

Table 1 Framework for factors that could affect the implementation, impact, and scaling up of the school resources

Domain	Factors and sources	Explanation
Teachers	Skills and competencies	Teacher's education and experience in relation to the lessons being taught.
	Understanding of the content being taught	Teacher's understanding of the content.
	Sufficient training	The extent to which the teachers received sufficient training in teaching the lessons
	Self-efficacy	Teacher's confidence in teaching the lessons.
	Fit to the teacher's teaching style and context (eg, class size)	Teachers' comfort with the instructions or ability to adapt the instructions to their style and context.
	Attitudes	Teachers' attitude towards new material (change), science, critical thinking and independent thinking by pupils (or their role as authorities in the classroom).
	Beliefs	Teachers' beliefs about the teaching methods or content (eg, what treatments work or the concepts).
	Emotions	Teachers' emotions, such as stress or anxiety.
	Motivation	Teachers' motivation to teach the material.
	Positive learning environment	Teachers' ability to create a positive learning environment; for example, encourage discussion, respond positively to questions, engage pupils.
Pupils	Literacy	Pupils' ablity to read and understand the material.
	Attendance	Pupils' attendance or reasons for poor attendance (eg, long distance to school or inability to pay school fees).
	Motivation to learn	Pupils' motivation to learn the new material.
	Attitudes	Pupils' attitudes towards learning, towards authorities, towards science, towards critical thinking.
	Beliefs	Pupils' beliefs about the content (eg, what treatments work or the concepts).
	Home environment	The extent to which the pupil's home environment encourages or discourages learning from the lessons.
	Differentiated instruction	The extent to which pupils different learning needs are met.
	Peer influence*	Positive or negative attitudes of other pupils towards the material.
Teaching materials	Value of the material	The extent to which the materials are valued by the teachers and pupils.
	Compatability with the curriculum	The extent to which the material fits with the rest of the curriculum and how it is taught.
	Appropriateness of the material*	The extent to which the materials are relevant, challenging and engaging.
	Credibility of the material*	The textent to which the teachers and pupils perceive the material as credible.

Continued

Table 1 Continued

Domain	Factors and sources	Explanation
School system and environment	Time constraints	The extent to which there is sufficient time to accommodate introducing the new material.
	Competing priorities*	The extent to which other priorities for the school, teachers or pupils limit introducing the material (eg, preparing for exams).
	School organisation and management*	The extent to which the school provides an environment that supports adoption of new subjects, material and teaching methods.
	School resources, particularly human resources*	The extent to which the school has adequate resources to introduce the new materials (eg, human resources, student/teacher ratio, teacher workload, classroom space and classroom resources, such as blackboards and acoustics).
	Attitudes and beliefs of head teacher and other teachers*	Attitudes or beliefs of colleagues that influence the teacher's interest in and ability to teach the material.
	Parent and community involvement*	Parents' attitudes towards the new material or how things are done at the school.
	Regulations*	Regulations (eg, Ministry of Education policies and regulations) that affect introducing the new material.
	Political environment *	Elements of the political environment that affect introducing the new material; for example, authoritarianism or teacher strikes.
	Bureaucracy*	Bureaucratic arrangements that delay or limit introduction of the new materials, or facilitate introducing them.
	Incentives and disincentives	Incentives or disincentives to introduce the new materials for teachers or head teachers.

*There were no key findings in relation to these factors or they were subsumed under another factor.

Table 2 Framework for potential adverse and beneficial effects

Potential adverse effects	Corresponding beneficial effects
Conflict between children and teachers due to children challenging their teachers	More open and engaging discussion of the basis of diverse claims or beliefs.
Conflict between children and parents due to children challenging their parents	Better understanding between children and parents due to children conversing with their parents about what they are learning and parents feeling more engaged with what their children are learning and engagement of parents in discussions of health issues.
Distrust of health professionals or conflict between children and health professionals*	Appropriate questioning of health professionals, better understanding and better healthcare.*
Conflict due to undermining of religious beliefs*	Engagement of children and others in discussion about religious beliefs and science.*
Shortened enjoyment of the innocence of childhood*	Increased enjoyment of school and childhood.
Nihilism or cynicism*	Healthy scepticism and appreciation science.
Other potential beneficial effects	
Impacts on teachers	The learning resources might improve the teachers' understanding and ability to apply the concepts being taught to the children.
Impacts on parents	The learning resources might indirectly improve parents' understanding and ability to apply the concepts being taught to the children.
Assertiveness	Children asking more questions and not taking things for granted.
Improved decision-making	Children making more thoughtful and informed decisions.
Nonviolent conflict resolution*	Claims being presented and addressed in a friendly manner even when there is a disagreement about the claim, as illustrated in the resources.
Friendship formation*	Friendly interactions between adults and children and among children, as illustrated in the resources.
Collaboration for problem solving*	Collaboration for problem solving among the children, as illustrated in the resources.
Creativity*	Thinking outside the box.
Numeracy	Improvements in numeracy, reflecting what is learnt in lessons 6 and 7 (on fair comparisons and the play of chance).

*There were no key findings in relation to these factors.

Data analysis

We used a framework thematic analysis approach, guided by the two frameworks described above (tables 1 and 2), and following the stages of familiarisation, coding, charting and interpretation of the data. We applied both of these frameworks to the same data set.

Two of the investigators (AN and DS) independently read and reread the transcripts. They then coded the data, using the factors included in the two frameworks, but also searching for additional factors. AN, DS, ADO, CG and SL then reviewed summaries of the coded data and considered additional factors suggested by the data. Disagreements were resolved by discussion. They also coded the qualitative data collected through the classroom observations and the teachers' evaluation forms. For each framework, the definitions and boundaries of each of the frameworks' factors were discussed among the investigators, and both frameworks were revised in line with the codes and categories that emerged from the data. We then charted the data by writing a summary of the findings for each framework factor. We then considered the extent to which the quantitative data from the teachers' evaluation forms[17] supported those findings and whether they suggested additional findings relevant for each framework that were not captured in our analysis of the qualitative data. Finally, using the summarised data, we explored the range and nature of phenomena, and possible explanations for the findings.

Quantitative data collection and analysis

All 85 teachers in the 60 schools in the intervention group of the trial completed a questionnaire at the end of the term during which the IHC lessons were taught.[17] They also completed a lesson evaluation form after each lesson.[17] The quantitative data we included from the lesson evaluation forms required teachers to rate suitability of the IHC materials on a Likert scale of 1–6 (1=lowest, 6=highest). Statistical analyses were performed with R (R Core Team, Vienna, Austria; V.3.4.3; using packages tidyverse and knitr).

Box 1 Data collection

Lesson evaluation forms: We included qualitative data from the lesson evaluation forms completed by teachers from the six schools selected for the process evaluation.[17] Data included teachers' suggestions for improving the informed health choices (IHC) materials specific to each lesson, what they liked and did not like and what facilitated their achievement or non-achievement of the intended lesson objectives. These were also entered into a spreadsheet.

Classroom observations: Each class in the 60 schools in the intervention group of the trial was observed at least once by a trained research assistant or one of the investigators (AN or DS) early on in the trial. In addition, we observed six classes a second time, after the teachers had become more familiar with using the resources. For observations, we used data collection forms to note how well the teachers adhered to the lesson plan, any problems that the teachers or children had with the lesson, and any aspects of the lesson that went particularly well.[17] The teaching was done in English and notes were taken in English, although teachers occasionally used a local language (predominantly Luganda). We noted the extent to which the children followed the lesson and participated actively. After the lesson, we collected the exercise books and recorded how well the children did on the exercises for the lesson that had been observed.

Interviews: We used individual interviews in order to gain an in-depth understanding of issues and further obtain detailed information about personal feelings, perceptions and opinions about the IHC intervention. Individual interviews took place in the participants' own environment, thus included children's schools' compounds, empty classrooms, empty teachers' staffrooms and policy-makers' personal offices at the district headquarters.

Interviews with district education officers, head teachers, and teachers: We conducted face-to-face interviews with 4 of the 5 district education officers, head teachers from 5 of the 6 schools and 10 grade 5 teachers who used the IHC primary school resources in the trial. We collected these data following completion of the intervention (which included nine lessons). We chose to only use face-to-face interviews to obtain in-depth data from district education officers and head teachers because it was not feasible to organise group discussions with them. We developed and used semistructured interview guides.[17] The guides focused primarily on questions related to barriers and facilitators (table 1), and strategies for scaling up use of the resources. We also included questions about potential adverse and beneficial effects (table 2). The interviewer included prompts for each of the domains and factors in the frameworks (tables 1 and 2), asking interviewees to reflect on their experiences and perceptions from their different perspectives. The adult interviews were scheduled to last an hour. A summary of each interview was provided to the interviewee for further comment.

Interviews with children: We used a semistructured interview guide for face-to-face interviews to elicit the views of children who used the IHC primary school resources in the trial.[17] We interviewed a total of 12 children, two from each of the six schools in the process evaluation. The children's interviews were scheduled to last no more than 30 min. The children were individually interviewed after we obtained parental consent and their assent. Although all the interviews took place on the school premises (classrooms, assembly halls and school compounds), in order to ensure confidentiality, the interviews were conducted in full view of a responsible adult (teacher) but not in close proximity to allow the children to freely share their experiences using the materials.

Focus group discussions: We used focus group discussions in order to gain detailed information about the different groups' feelings, perceptions and opinions.

Continued

Box 1 Continued

Focus group discussions with children and teachers: We used focus group discussions with children and teachers to explore barriers and facilitators to using the resources, as well as potential adverse and beneficial effects. We carried out three focus group discussions with children and two with teachers. The focus group discussions included six to eight participants, with clear ground rules (including confidentiality) agreed in advance. The children's focus group discussions were scheduled to last 45 min while the adults (teachers) focus group discussions lasted 1.5 hours. Each group was moderated by AN or DS using a guide[17] and assisted by an observer who took notes. We used an iterative process to develop focus group guides drawing on issues emerging from the initial individual interviews to revise the questions and to create prompts for the discussions. We conducted further interviews after we had conducted some focus group discussions and had some preliminary findings from these. This allowed us to explore issues for which more detailed data were needed, or to capture the views of particular subgroups (eg, poorer children or children who were not doing well in class).

Interviews with parents: We used a semistructured interview guide for the individual face-to-face interviews with parents whose children used the IHC primary school resources in the trial.[17] This interview guide included questions about the parents' perspectives on how the resources were used, barriers and facilitators to their use, potential adverse and beneficial effects, and potential effects of the school resources on parents. We interviewed a total of 13 parents of children who used the IHC primary school resources in the trial. The parents were selected from participants in the podcast trial. We included both parents who listened to the IHC podcast series and parents in the comparison group of the podcast trial, who did not listen to the podcast series, in order to gain a balanced view of opinions.

Interviews with the lead investigators: AN and DS were responsible for implementing the intervention in the field. Given the importance of their role in the trial and the process evaluation, two other investigators (SL and CG) in turn interviewed AN and DS to capture their thoughts and experiences and their reflections on the findings from the other data sources.

Appraisal of the certainty of the findings of the process evaluation

We assessed the certainty of the findings using a modified version of the GRADE-CERQual ('Confidence in the Evidence from Reviews of Qualitative research') approach[25 26] (see online supplementary file 2). Although CERQual has been designed for findings emerging from qualitative evidence syntheses, several components of the approach are suitable for findings based on multiple primary sources of qualitative data. So far as we are aware, this is the first time that a modified version of GRADE-CERQual has been used to assess findings from a single study rather than from an evidence synthesis.

Integration of the findings of the process evaluation with the findings of the trial

We used a logic model approach to organise the findings of this process evaluation with the findings of the trial. First AN and DS organised the findings into chains of events that may have led to the outcomes of the trial and additional outcomes that we explored (potential adverse and beneficial effects). Findings and outcome measures were categorised as follows in the logic model:

- Attributes of the intervention.
- Effect modifiers.
- Intermediate outcomes.
- Observed and potential effects.

After discussion, the investigators revised the logic model iteratively until there was agreement on a final model.

Patient and public involvement statement

Primary school children, parents, teachers, headteachers and policy-makers (district educational officers) participated in providing structured feedback. They were not otherwise involved in the design of the process evaluation or the analysis. This study was preceded by the prioritisation of key concepts[9] and the development of the IHC resources.[10] Those processes engaged key stakeholders (teachers, children and policy-makers) in the prioritisation of key concepts for inclusion in the resources, brainstorming workshops to generate ideas for the resources, and consultation workshops to discuss the feedback gathered during user-testing and piloting of earlier versions of the resources.

RESULTS

The main findings are summarised in online supplementary file 2, and we have integrated the findings into a logic model in figure 2. Quantitative data are summarised descriptively in online supplementary file 3. We have used the logic model to organise the results.

The intervention

Value of the IHC intervention

Most of the children, teachers and parents liked the IHC materials because they found them beneficial, interesting and fun for the children. Teachers valued that the IHC content addressed both social and academic issues.

Compatibility with the curriculum

All the teachers that taught the IHC lessons emphasised that the IHC content was important that it should be added to the curriculum, and that it was compatible with the primary school curriculum for science. We found strong support for this finding in the quantitative data in the teachers' end-of-term assessment: 99% thought that what the children learnt was very important or important; and 95% agreed or strongly agreed that they liked the content of the lessons.[14]

Figure 2 Logic model. IHC, informed health choices.

 Nsangi A, *et al. BMJ Open* 2019;**9**:e030787. doi:10.1136/bmjopen-2019-030787

Compatibility with teaching styles

Most of the teachers found that the design of the IHC lessons was compatible with their teaching styles, especially the use of multiple examples in the teachers' guide. Several teachers, however, felt that specific aspects of the IHC lessons were in conflict with their teaching styles, such as summarising the content of each lesson. These were not perceived as major challenges.

Differentiated instruction

The teachers felt that the Teachers' Guide[27] allowed sufficient flexibility for teachers to employ different ways of doing things to accommodate different teaching styles. Observations and video recordings from the trial document different teaching styles used to teach the IHC lessons. Examples of the videos can be found here. In the end-of-term teachers' assessment, 90% agreed or strongly agreed that the instructions for how to teach the lessons fit with their teaching style; 92% agreed or strongly agreed that they liked the way the teaching materials were organised; and 88% agreed or strongly agreed that they adapted the instructions to fit their teaching style.

Teachers pointed out that the way the materials were structured enabled children of different capabilities to participate in the IHC lessons more than they would participate in other lessons. However, several teachers found the material challenging for children who were not fluent in English or who had poor reading skills. This is consistent with the finding in the trial that the effect of using the IHC primary school resources was larger for children with better reading skills.[14]

Training and understanding of the content being taught

The introductory workshop that was part of the IHC intervention was seen as appropriately conducted by the teachers with many crediting it for the smooth implementation of the project. The majority of the teachers acknowledged that the IHC content was new to them, and some expressed concern about their understanding of the IHC content. All participants thought that the introductory workshop was too short. Despite these concerns, 94% of the teachers agreed or strongly agreed in their end-of-term assessment that they understood the content of the lessons and 97% responded that they learnt very much or much. This is further supported by the proportion of teachers who had mastered the material by the end of the term (72%) compared with the teachers in the control schools (15%).[14]

Adding on to the curriculum

The IHC lessons were viewed as an add on to what was already in the curriculum. Several teachers emphasised the importance of incorporating the IHC lessons into an already a packed primary school curriculum. This was also a cause of concern for the district education officers, and one of the main reasons why some schools declined to participate in the study.[14]

Effect modifiers

Incentives

Several teachers identified incentives that motivated them to teach the IHC lessons. These included having head teachers and school owners that were supportive, ongoing support from the research team and simply having enough textbooks.

Teachers' competencies

Teachers said they had diverse competencies that were important for delivering the IHC content to the children, such as communication and teneral teaching skills. Most felt that it was important that they were science teachers.

Positive learning environment

The majority of the teachers shared ways in which they created a positive learning environment for the children during the IHC lessons, including use of role playing, using relevant examples and allowing children to express their opinions.

Teachers' beliefs

Many teachers had beliefs that were in conflict with some of the examples and sometimes directly in conflict with a key concept, particularly the concept that widely used treatments or treatments that have been used for a long time are not necessarily effective or safe. Although many teachers expressed concerns about it being difficult to reconcile their beliefs with the IHC content, the quantitative data suggest that this may not have had an important impact on their ability to apply the concept or teach it.

Children's beliefs

Children were less likely to identify conflicts between their beliefs and the IHC lessons than the teachers were. However, the children we interviewed seemed to struggle with the key concept that personal experiences or anecdotes (stories) are an unreliable basis for assessing the effects of most treatments. Nonetheless, 30% more children in the intervention schools answered both questions on this concept correctly compared with children in the control schools.[15]

Intermediate effects

Teachers' motivation

The majority of the teachers felt motivated while teaching the IHC lessons because they felt that the content was important for the children and for themselves. Others felt that the way the IHC programme was introduced to them at the introductory workshop motivated them. Others felt that the support provided by the IHC team and their head teachers motivated them. This included provision of adequate materials, carrying out observational visits, progress calls and encouragement.

Teachers' self-efficacy

While the content was new for most teachers, this affected their confidence differently. Some felt that their professional training equipped them to teach new material. Many teachers also noted that the children's positive response and enthusiasm about the lessons made them more confident. However, others initially lacked confidence. A few teachers felt uncertain as a result of teaching the lessons, particularly when they felt put on the spot by children asking them questions. Nevertheless, by the end of the term, 95% of the teachers strongly agreed or agreed that they were confident about their ability to teach the lessons, and 94% strongly agreed or agreed that they liked teaching something new. In the lesson assessments, most teachers indicated that it was easy for them to teach the lessons and that they were comfortable teaching the lessons.

Teacher's attitudes

Teachers mostly had positive attitudes towards the IHC lessons, particularly in relation to the content being new and valuable to them as well as to the children. Several of them vividly expressed how they felt during the lessons. One noted: 'It was something I can't even describe, being a new idea brought to me, I loved it so much. I was teaching learners, as I was also teaching myself.'

There was strong support for this from the quantitative data: 98% liked very much or liked teaching the lessons; 99% thought that what the children learnt was very important or important; 94% thought that the children learnt very much or much from the lessons and 98% strongly agreed or agreed that it is important to teach children to think critically.

Children's motivation to learn

The children that attended the IHC lessons were motivated. They enjoyed the lessons and looked forward to them. They liked both the design and the content of the book, including the pictures, the characters, the games and being able to colour in the pictures in their exercise books.

Positive learning environment

Most of the children indicated that they experienced the IHC lessons positively. They found the lessons enjoyable and the books interesting. Teachers noted that children were more active during the IHC lessons compared with their normal lessons, that they were enthusiastic, and that they were eager to attend. The majority of the parents also mentioned that their children enjoyed the IHC lessons, noting that they were reading their books and sharing what they had learnt with people at home. This is consistent with the trial results. When asked 'How much did you like what you learned as part of the lessons with The Health Choices Book?', 78% of children in the intervention group responded that they 'liked the lessons very much' and another 16% responded that they 'liked the lessons'.

Time constraints

Although the majority of the teachers were able to complete all the nine IHC lessons, this was not always to their satisfaction due to other competing priorities. The quantitative data from the lesson evaluation forms showed that the amount of time different teachers used preparing and teaching lessons was similar across lessons, but varied among teachers. Most teachers reported using between 5 and 30 min preparing for each lesson and between 45 and 90 min teaching each lesson. Most teachers felt that they spent close to the right amount of time preparing and teaching each lesson, but some felt that they spent too much or too little time preparing and teaching each of the lessons. There was little correlation between the amount of time a teacher used preparing and teaching each lesson and whether they thought that was too much or too little time (see online supplementary file 3).

Attendance

Teachers pointed out that absenteeism was a common problem. The children whose parents had paid the school fees on time said that they attended all the lessons. In the trial, 10% of the children in the intervention schools and 29% in the control schools did not complete the test. To a large extent, absenteeism was attributed to parents' failure to pay their children's school fees in time.

Scaling up

Parent and community involvement

Teachers, parents and district education authorities emphasised the importance of involving other stakeholders, including parents and the communities at large, for the lessons to be effective.

Collaboration with policy-makers

Several teachers emphasised the importance of working in collaboration with the Ministry of Education and the National Curriculum Development Centre to ensure the IHC lessons are incorporated into the primary school curriculum. The majority of the teachers, parents and children interviewed supported spreading the IHC programme to other schools and other age groups, including infant, middle and upper primary sections.

Potential beneficial and adverse effects

Beneficial effects

Nearly everyone interviewed thought that both the children and teachers learnt important lessons from the IHC resources and many felt that the lessons improved their decision-making. Several parents observed that children gained confidence and started asking important questions about the benefits and harms of treatments before decisions were made. Teachers also noted impacts on their own learning and decision-making. A teacher gave this example: 'Somebody came to school and asked us to buy some food supplements. Then I asked myself: 'Are these foods really well researched?' In fact, I did not buy (them).'

Some teachers and children also noted beneficial impacts on English and numeracy skills.

Adverse effects

Although teachers found the IHC lessons enjoyable, some reported having experienced stress because of teaching something new and it being additional to their usual subjects.

The majority of teachers and parents expressed concern about the potential conflicts between themselves and the children resulting from children sometimes challenging their authority such as asking questions or refusing to take instructions from those in authority. However, there were no reports of actual conflicts.

DISCUSSION

Facilitating factors

The findings of this process evaluation are consistent with the findings of the trial, in that most of the factors that were identified were facilitators rather than barriers to the implementation of the intervention (figure 2). The findings suggest that children, teachers and parents appreciated the IHC school intervention. Teachers found the IHC lessons compatible with the curriculum and their teaching styles; and the materials enabled teachers to apply differentiated instruction. Effect modifiers included teachers' skills and competencies and positive learning environments. These contributed to intermediate effects, including teachers' motivation, self-efficacy, positive attitudes and a positive overall experience, which in turn contributed to the IHC resources having a large effect on the ability of the children to assess claims about treatment effects.

Children, teachers, parents, head teachers and district education officers valued the IHC primary school resources and their content. We believe this is due, in large part, to the fact that we spent 3 years developing them using a human-centred design approach.[11]

Through this approach we tried out many ideas, developed some into prototypes and gradually improved these in cycles, based on extensive feedback from users and through observation of use in classrooms. This enabled us to identify and resolve problems early on, leading to solutions that people valued, for example, use of a comic book format, creating characters that appealed to the children, adding local language vocabulary, building in activities teachers could carry out in large classes without extra materials and adding examples that were familiar in the East African context.

The IHC primary school resources were designed to support teachers who were not familiar with the content, and the workshop was designed to introduce the project in general (goals and expectations), to answer questions and to offer clarifications. It was not designed to teach the content to the teachers. Although some teachers were concerned about the duration of the training workshop, most of the teachers in the intervention group mastered the IHC key concepts after teaching them. This is likely attributable, at least in part, to their having learnt the IHC key concepts as they taught the children. This suggests that, although some teachers perceived the need for a longer workshop, this might not be necessary either for them to master the IHC key concepts or for them to teach them effectively. It might be helpful to inform teachers about this and to reassure them.

Support from the school authorities and from colleagues played an important role in ensuring that the teachers were allowed adequate time to teach the IHC lessons. This likely contributed to the effectiveness of the intervention in the trial.[14] Effectively scaling up use of the IHC lessons will undoubtedly require the support of school authorities to ensure that teachers have sufficient time.

The children enjoyed the lessons and looked forward to them. Their positive attitudes towards the materials and the lessons likely played an important role in the effect that the intervention had on their ability to assess treatment claims. Many teachers noted that the children's positive response and enthusiasm about the lessons made them more confident.

The IHC materials facilitated switching from English to local languages. This made it possible for children to ask questions in the language with which they were most comfortable. This is not common for subjects in the normal upper primary school curriculum in Uganda. Children clearly stated that they generally appreciated being able to use more than one language during the IHC lessons. Several studies have noted the benefits accrued over time on functional literacy (reading and writing) when children are able to study in their mother tongue.[28–30]

Impeding factors

Many teachers identified conflicts between their beliefs and personal practices and the IHC content. These were particularly in relation to herbal remedies and the concept that widely used treatments or those that have been in use for a long time are not necessarily beneficial or safe. Paying particular attention to examples and IHC key concepts about which the teachers may have conflicting beliefs might help to address this. For example, it might help to acknowledge in the Teachers' Guide[27] and the training workshop that teachers and others commonly believe that treatments are effective when, in fact, their effects are uncertain or there is evidence to the contrary. Open discussion about disagreements, the logic and the evidence underlying relevant IHC key concepts; and use of compelling examples that do not directly challenge the teachers' beliefs might be helpful. For example, ineffective interventions that were widely used for decades or centuries, which are no longer used, such as bloodletting, could be used as examples.

Children's beliefs may be less resistant to change than adults' beliefs,[31] which may be reflected in the fact that children were less likely to identify conflicts between their beliefs and the IHC lessons than the teachers were. On the other hand, some children struggled with the key concept that personal experiences or anecdotes are an unreliable basis for assessing the effects of most treatments. It also might be helpful for this key concept to provide teachers and children with a variety of examples, including some that are less likely to challenge their prior beliefs about the effects of specific treatments.

Absenteeism, likely attributable to a large extent to the parent's failure to pay tuition fees, was another barrier identified by the teachers and the lead investigators. This is a systemic problem, not specific to the IHC lessons.[32] Another important barrier, which was not identified in the process evaluation, but is undoubtedly the biggest challenge, is the cost of the intervention.[14]

In the context of our trial, we were not able to remove anything from teachers' already heavy workloads. If the IHC lessons were incorporated into the curriculum, rather than taught as an add-on, as in the trial, that might reduce the burden on teachers and increase the effectiveness of the intervention. A second reason why the intervention might be even more effective when scaled up, would be that teachers would not be teaching it for the first time after the first year and the material would no longer be new to them. On the other hand, they would not have support from the research team. Some teachers suggested that this might have facilitated implementation of the intervention, even though the research team only observed lessons when they visited schools and did not provide feedback or help. In addition, teachers indicated that the training workshop, which was taught by the two principal investigators was important. It is uncertain whether it would be feasible to offer a comparable workshop to all of the teachers in the country who would be responsible for teaching the IHC lessons.

Potential effects

The children shared what they learnt with their families. Although having a child in a school that used the primary school resources had little, if any, short-term effect on the parents' test scores for parents who participated in the podcast trial,[16] after 1 year the mean score of parents with a child in an intervention school was 4% higher than that of parents with a child in a control school and 12% more parents had a passing score.[33] This finding is promising in terms of the potential for the intervention to benefit families of the children and not just the children themselves.

Other potential benefits that were identified, besides the value of what was learnt, include improved decision-making by teachers, as well as by children, improvements in English and numeracy, and improved relationships between children and adults.

The most common concern was the risk of conflict between children and authorities, including teachers, parents and healthcare professionals. The concern was that children questioning claims made by authorities might be viewed as disrespectful and undermining of their authority. Undermining of religious and cultural beliefs was also identified as a potential adverse effect. There were incidents of children challenging authorities, including teachers, parents, head teachers and members of the research team. However, none of the participants in the trial or the process evaluation reported actual conflicts and none were observed.

Strength and limitations

An important limitation of this study is that the investigators were responsible for both developing and evaluating the intervention. This could have led us to emphasise participants' positive experiences of the intervention when designing the process evaluation and when collecting and analysing the data. In addition, as most respondents were aware that the lead investigators were responsible for the intervention itself, there may have been a desire to please the investigators by giving positive reports.[34] Although we tried to address this by making it clear to the respondents that the IHC materials were being tested, not them, there may have been some desirability bias.[35]

An important strength of the study is our use of data collection triangulation. We gathered data using a variety of methods, and explored whether the findings from different sources challenged or supported each other. In addition, based on a modified CERQual approach, we have high confidence in most of our findings. Exceptions are our findings regarding children's beliefs, time constraints, incentives and disincentives (see online supplementary file 2).

CONCLUSIONS

The extent to which children, teachers, parents, head teachers and district education officers valued the IHC primary school resources played a key role in facilitating the success of our intervention. The key barrier that we identified was that the IHC lessons were an add-on, rather than being incorporated in the national curriculum.

Author affiliations
[1]Department of Medicine, Makerere University College of Health Sciences, Kampala, Uganda
[2]Institute of Health and Society, University of Oslo Faculty of Medicine, Oslo, Norway
[3]Centre for Informed Health Choices, Norwegian Institute of Public Health, Oslo, Norway
[4]Health Systems Research Unit, South African Medical Research Council, Tygerberg, South Africa
[5]Regional Centre for Child and Adolescent Mental Health (Eastern and Southern Norway), Oslo, Norway
[6]Department of Public Health, University of Rwanda, Butare, Rwanda
[7]Tropical Institute of Community Health and Development, Kisumu, Kenya

Acknowledgements We would like to especially thank the participating schools, pupils, teachers, parents and educational authorities in this process evaluation. Special thanks to our team of dedicated research assistants namely; Esther Nakyejwe, Hilda Mwebaza, Rita Tukahirwa, Nora Rosenbaum, David Ssimbwa, Peter Lukwata, Solomon Segawa and Martin Mutyaba.

Contributors AN and DS were responsible for data collection in Uganda. AN, DS, ADO, CG and SL participated in the analyses of the qualitative data. AN, DS, ADO, CG and SL drafted the first draft of the manuscript. All the other authors (MO, SR, AD, LN, MK, AF and NKS) except for CJR reviewed the protocol, provided input on the manuscript, and agreed on the final version. CJR performed the statistical analyses. ADO, NKS and AF had the primary responsibility for overseeing the study.

Funding The Research Council of Norway, project number 220603/H10, funded the study. The Norwegian Institute of Public Health, recipient of the grant from the Research Council of Norway, is the coordinating centre for the Informed Health Choices project. This work was also partially supported by a Career Development Award from the DELTAS Africa Initiative grant # DEL-15-011 to THRiVE-2. The DELTAS Africa Initiative is an independent funding scheme of the African Academy of Sciences (AAS)'s Alliance for Accelerating Excellence in Science in Africa (AESA) and supported by the New Partnership for Africa's Development Planning and Coordinating Agency (NEPAD Agency) with funding from the Wellcome Trust grant # 107742/Z/15/Z and the UK government.

Disclaimer The views expressed in this publication are those of the author(s) and not necessarily those of AAS, NEPAD Agency, Wellcome Trust or the UK government.

Competing interests None declared.

Patient consent for publication Obtained.

Ethics approval The study was approved by Makerere University Institutional Review Board and the Uganda National Council of Science and Technology as part of the Supporting Informed Healthcare Choices in Low-income Countries Project (grant no. ES498037).

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available on reasonable request.

REFERENCES

1. Lewis M, Orrock P, Myers S. Uncritical reverence in cm reporting: assessing the scientific quality of Australian news media reports. *Health Sociol Rev* 2010;19:57–72.
2. Moynihan R, Bero L, Ross-Degnan D, *et al*. Coverage by the news media of the benefits and risks of medications. *N Engl J Med Overseas Ed* 2000;342:1645–50.
3. Zarcadoolas C, Pleasant A, Greer D. *Advancing health literacy: a framework for understanding and action*. 368. 1st edn. San Francisco, CA: Jossey-Bass, 2006.
4. Austvoll-Dahlgren A, Oxman AD, Chalmers I, *et al*. Key concepts that people need to understand to assess claims about treatment effects. *J Evid Based Med* 2015;8:112–25.
5. Weinstein ND. What does it mean to understand a risk? Evaluating risk comprehension. *JNCI Monographs* 1999;1999:15–20.
6. Lloyd AJ. The extent of patients' understanding of the risk of treatments. *Qual Health Care* 2001;10 Suppl 1(Suppl 1):i14–18.
7. Austvoll-Dahlgren A, Nsangi A, Semakula D. Interventions and assessment tools addressing key concepts people need to know to appraise claims about treatment effects: a systematic mapping review. *Syst Rev* 2016;5.
8. Austvoll-Dahlgren A, Guttersrud Øystein, Nsangi A, *et al*. Measuring ability to assess claims about treatment effects: a latent trait analysis of items from the 'Claim Evaluation Tools' database using Rasch modelling. *BMJ Open* 2017;7:e013185.
9. Nsangi A, Semakula D, Oxman AD, *et al*. Teaching children in low-income countries to assess claims about treatment effects: prioritization of key concepts. *J Evid Based Med* 2015;8:173–80.
10. Nsangi A, Semakula D, Rosenbaum S, *et al*. Development of the informed health choices resources in four countries to teach primary school children to assess claims about treatment effects. *Pilot Feasibility Stud* 2019. Submitted.
11. IDEO. *Human centered design toolkit*. 1st edn. San Francisco: IDEO, 2011.
12. Semakula D, Nsangi A, Oxman M, *et al*. Development of mass media resources to improve the ability of parents of primary school children in Uganda to assess the trustworthiness of claims about the benefits and harms of treatments. *Pilot Feasibility Stud* 2019. Submitted.
13. Sandoval WA, Sodian B, Koerber S, *et al*. Developing children's early competencies to engage with science. *Educ Psychol* 2014;49:139–52.
14. Nsangi A, Semakula D, Oxman AD, *et al*. Effects of the informed health choices primary school intervention on the ability of children in Uganda to assess the reliability of claims about treatment effects: a cluster-randomised controlled trial. *Lancet* 2017;390:374–88.
15. Nsangi A, Semakula D, Oxman AD, *et al*. Effects of the informed health choices primary school intervention on the ability of children in Uganda to assess the reliability of claims about treatment effects after one year: a cluster-randomised trial. *Trials* 2019. Submitted.
16. Semakula D, Nsangi A, Oxman AD, *et al*. Effects of the informed health choices podcast on the ability of parents of primary school children in Uganda to assess claims about treatment effects: a randomised controlled trial. *Lancet* 2017;390:389–98.
17. Nsangi A, Semakula D, Glenton C, *et al*. Resources to teach primary school children in low-income countries to assess claims about treatment effects; Protocol for a process evaluation. *IHC working paper*, 2016. Available: http://www.informedhealthchoices.org/wp-content/uploads/2016/08/Claim-cut-off-IHC-Working-Paper-2017-01-09.pdf [Accessed 30 Oct 2018].
18. Tweed A. Designing effective science instruction (what works in science classrooms?). *National Science Teachers Association* 2009.
19. Klingner JK, Ahwee S, Pilonieta P, *et al*. Barriers and facilitators in scaling up Research-Based practices. *Except Child* 2003;69:411–29.
20. Forman SG, Olin SS, Hoagwood KE, *et al*. Evidence-based interventions in schools: developers' views of implementation barriers and facilitators. *School Ment Health* 2009;1:26–36.
21. Ryan P. Integrated theory of health behavior change. *Clinical Nurse Specialist* 2009;23:161–70.
22. Ayres BJ, Meyer LH, Erevelles N, *et al*. Easy for you to say: teacher perspectives on implementing most promising practices. *Journal of the Association for Persons with Severe Handicaps* 1994;19:84–93.
23. Flottorp SA, Oxman AD, Krause J, *et al*. A systematic review and synthesis of frameworks and taxanomies of factors that prevent and enable improvements in healthcare professional practice. *Implement Sci* 2013;8.
24. Carlsen B, Glenton C. What about N? A methodological study of sample-size reporting in focus group studies. *BMC Med Res Methodol* 2011;11.
25. Lewin S, Glenton C, Munthe-Kaas H, *et al*. Using qualitative evidence in decision making for health and social interventions: an approach to assess confidence in findings from qualitative evidence syntheses (GRADE-CERQual). *PLoS Med* 2015;12:e1001895.
26. Lewin S, Booth A, Glenton C, *et al*. Applying GRADE-CERQual to qualitative evidence synthesis findings: introduction to the series. *Implement Sci* 2018;13.
27. Informed Health Choices Group. *Teachers' Guide for The Health Choices Book: Learning to think carefully about treatments. A health science book for primary school children*. Oslo: Norwegian Institute of Public Health, 2016.
28. Cook V. Using the first language in the classroom. *Canadian Modern Language Review* 2001;57:402–23.
29. Costley T. English as an additional language, policy and the teaching and learning of English in England. *Language and Education* 2013;28:278–92.
30. Alshammari MM. The use of the mother tongue in Saudi EFL classrooms. *JIER* 2011;7:95–102.
31. Duschl RA, Schweingruber A, Shouse AW, eds. *Committee on science learning, kindergarten through eighth grade. How children learn science. taking science to school: learning and teaching science in grades K-8*. Washington, DC: The National Academies Press, 2007.
32. Tamusuza A. Leaving school early: the quest for universal primary education in Uganda. *J Statisque Africain* 2011;13:110–51. https://www.afdb.org/fileadmin/uploads/afdb/Documents/Publications/Leaving%20School%20Early%20The%20Quest.pdf
33. Semakula D, Nsangi A, Oxman AD, *et al*. Effects of the informed health choices podcast on the ability of parents of primary school children in Uganda to assess the trustworthiness of claims about treatment effects, one-year follow-up: a randomised trial. *Trials* 2019. Submitted.
34. Knowles ES, Nathan KT, Kobi TN. Acquiescent responding in Self-Reports: cognitive style or social concern? *J Res Pers* 1997;31:293–301.
35. King MF, Bruner GC. Social desirability bias: a neglected aspect of validity testing. *Psychol Mark* 2000;17:79–103.

Additional file 3. Quantitative analyses for school process evaluation

Table 1. Teachers' assessments of the lessons at the end of the term

Question	Responses (N=85)				
How important do you think fair comparisons of educational interventions are?	Missing 6 (7.1%)	Very unimportant 5 (5.9%)	Unimportant 0 (0%)	Important 19 (22%)	Very important 55 (65%)
How much did you like teaching the lessons?	Missing 1 (1.2%)	Not at all 0 (0%)	Did not like 1 (1.2%)	Liked 32 (38%)	Lived very much 51 (60%)
How important is what the children learned from the lessons?	Missing 1 (1.2%)	Unimportant 0 (0%)	Little importance 0 (0%)	Important 18 (21%)	Very important 66 (78%)
How easy or difficult was it for the children in your class to understand the lessons?	Missing 1 (1.2%)	Very difficult 0 (0%)	Difficult 9 (11%)	Easy 71 (84%)	Very easy 4 (4.7%)
How much do you think the children learned from the lessons?	Missing 1 (1.2%)	Very little 0 (0%)	Little 4 (4.7%)	Much 55 (65%)	Very much 25 (29%)
How much do you think you learned from the lessons?	Missing 1 (1.2%)	Very little 0 (0%)	Little 1 (1.2%)	Much 18 (21%)	Very much 65 (76%)

Nsangi A, *et al*. *BMJ Open* 2019; 9:e030787. doi: 10.1136/bmjopen-2019-030787

Table 2. Teachers' assessments of the lessons at the end of the term - continued

Statement	Missing	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
The training that I received on how to teach the lessons was necessary for me to be able to teach the lessons well	1 (1.2%)	0 (0%)	1 (1.2%)	1 (1.2%)	34 (40%)	48 (56%)
The training that I received on how to teach the lessons was	2 (2.4%)	0 (0%)	10 (12%)	5 (5.9%)	45 (53%)	23 (27%)
I understood the content of the lessons well	1 (1.2%)	0 (0%)	0 (0%)	4 (4.7%)	42 (49%)	38 (45%)
I am confident about my ability to teach the lessons to the children	1 (1.2%)	0 (0%)	0 (0%)	3 (3.5%)	28 (33%)	53 (62%)
The teacher's guide was useful in conducting the lessons	1 (1.2%)	0 (0%)	0 (0%)	0 (0%)	25 (29%)	59 (69%)
The instructions for how to teach the lessons fit my teaching style well	1 (1.2%)	0 (0%)	4 (4.7%)	4 (4.7%)	54 (64%)	22 (26%)
The lesson activities helped children understand the content better	1 (1.2%)	0 (0%)	0 (0%)	4 (4.7%)	40 (47%)	40 (47%)
I had to adapt the instructions to fit my teaching style	2 (2.4%)	0 (0%)	2 (2.4%)	6 (7.1%)	56 (66%)	19 (22%)
I liked teaching something new	1 (1.2%)	1 (1.2%)	1 (1.2%)	2 (2.4%)	36 (42%)	44 (52%)
I think it is important to teach children to think critically	1 (1.2%)	0 (0%)	0 (0%)	1 (1.2%)	16 (19%)	67 (79%)
I do not like being challenged by children in my class	3 (3.5%)	20 (24%)	26 (31%)	4 (4.7%)	15 (18%)	17 (20%)
I like the content of the lessons	1 (1.2%)	0 (0%)	2 (2.4%)	1 (1.2%)	50 (59%)	31 (36%)
I liked the way the teaching materials and lessons were organised	2 (2.4%)	0 (0%)	1 (1.2%)	4 (4.7%)	51 (60%)	27 (32%)

Table 3. Correlations (and p-values) between time spent preparing or class time and teachers' assessments of whether that was too little or too much time

Lesson	Preparation	p-value	Class	p-value
1	0.2274786	0.07	0.5644887	<0.0001
2	0.1978938	0.12	0.1508329	0.22
3	0.1278681	0.31	0.2499880	0.04
4	-0.0066538	0.95	0.1039663	0.36
5	0.0569024	0.64	0.2783984	0.02
6	0.0238236	0.85	0.2370305	0.05
7	0.0819400	0.52	0.0137939	0.91
8	0.0880126	0.49	0.1437234	0.24
9	0.1700126	0.23	0.2679309	0.04

All correlations are Spearman (non-parametric) coefficients.

Table 4. Distributions of teachers' assessments of suitability of materials, summarized by 25th, 50th (median), and 75th percentiles

Material	Lesson	25%	50%	75%
Children's Book	1	4	4	6
	2	4	5	6
	3	4	4	6
	4	4	4	6
	5	4	4	6
	6	4	4	6
	7	4	4	6
	8	4	4	6
	9	4	4.5	6
Teacher's Guide	1	4	5	6
	2	4	5	6
	3	4	4	6
	4	4	5	6
	5	4	5	6
	6	4	4	6
	7	4	4	6
	8	4	5	6
	9	4	4	6
Activity	1	4	4	6
	2	4	4	6
	3	4	4	6
	4	4	4	6
	5	4	5	5
	6	4	4	5
	7	4	4	6
	8	4	4	5
	9	4	4	6
Exercises	1	4	4	5
	2	4	5	6
	3	4	4	5
	4	4	4	5
	5	4	4	6
	6	4	4	5
	7	4	4	6
	8	4	4	5
	9	4	4	6

Nsangi A, *et al*. *BMJ Open* 2019; 9:e030787. doi: 10.1136/bmjopen-2019-030787

Table 5. Distributions of teachers' assessments of children's success, summarized by 25th, 50th (median), and 75th percentiles

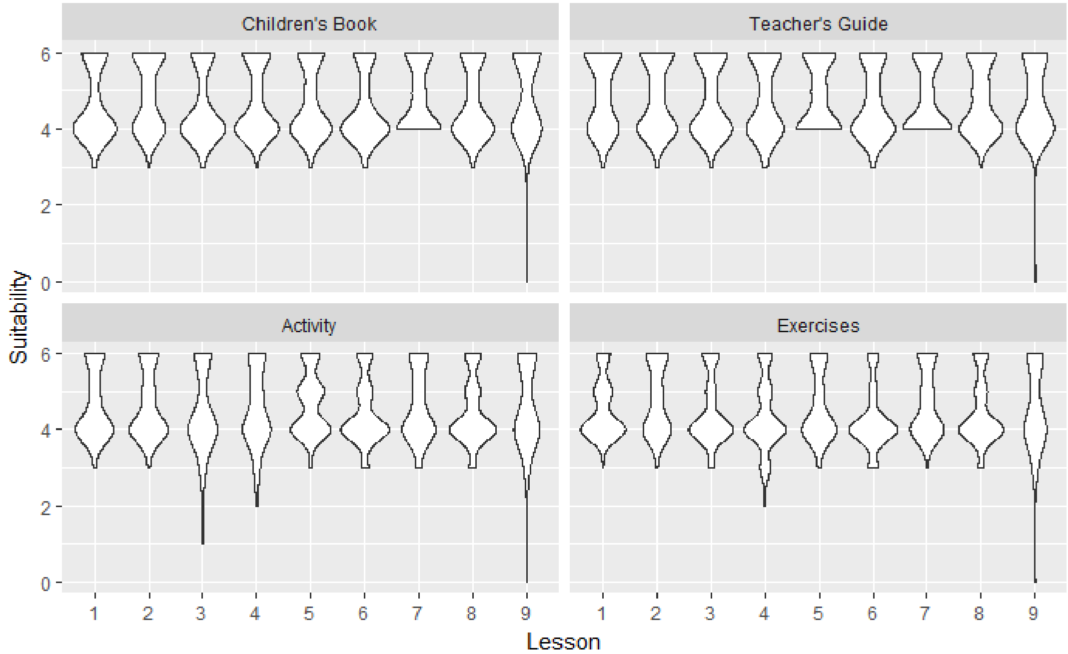
Question	Lesson	25th	50th	75th
Appropriateness of objectives	1	4	6	6
	2	4	6	6
	3	4	6	6
	4	4	6	6
	5	5	6	6
	6	4	5	6
	7	4	6	6
	8	5	6	6
	9	4	6	6
Difficulty for children	1	4	4	4
	2	4	4	5
	3	3	4	5
	4	4	4	5
	5	4	4	5
	6	4	4	4
	7	4	4	5
	8	4	4	5
	9	4	4	6
Difficulty for teacher	1	4	4	6
	2	4	4	6
	3	4	4	5
	4	4	4	6
	5	4	4	5.5
	6	4	4	5
	7	4	4	6
	8	4	4	5
	9	4	5	6
Comfort teaching the lesson	1	4	4.5	6
	2	4	4	6
	3	4	4	6
	4	4	4	6
	5	4	4	6
	6	4	4	5
	7	4	4	6
	8	4	4	6
	9	4	5	6
Too much or little time preparing	1	3	3	4
	2	3	3	3
	3	3	3	3
	4	3	3	3
	5	3	3	3
	6	3	3	3
	7	3	3	3
	8	3	3	3
	9	3	3	3
Too much or little time teaching	1	3	3	4
	2	3	3	3
	3	3	3	3
	4	3	3	3
	5	3	3	3
	6	3	3	3
	7	3	3	3
	8	3	3	3
	9	3	3	3
Interestingness for children	1	4	5	6
	2	4	5.5	6
	3	4	4	6
	4	4	4	6
	5	4	4	6
	6	4	4	6
	7	4	4	6
	8	4	4	6
	9	4	4	6

Table 6. Distributions of preparation and class time used (in minutes), summarized by 25th, 50th (median), and 75th percentiles

Question	Lesson	25%	50%	75%
Preparation time	1	15.00	20	35.00
	2	18.75	20	30.00
	3	20.00	30	40.00
	4	15.00	20	30.00
	5	20.00	20	40.00
	6	20.00	20	35.00
	7	18.00	20	30.00
	8	10.00	20	30.00
	9	15.00	20	30.00
Class time	1	60.00	80	90.00
	2	60.00	80	86.25
	3	60.00	79	90.00
	4	60.00	80	80.00
	5	60.00	80	83.75
	6	70.00	80	90.00
	7	60.00	80	81.25
	8	60.00	80	80.00
	9	60.00	80	86.25

Nsangi A, *et al*. *BMJ Open* 2019; 9:e030787. doi: 10.1136/bmjopen-2019-030787

Figure 1. Teachers' assessments of material suitability by lesson



1. Rate the suitability of the *children's book* for teaching and learning this lesson for your class

0	1	2	3	4	5	6
Very bad		Bad		Good		Very good

2. Rate the suitability of the *teacher's guide* for teaching and learning this lesson to your class

0	1	2	3	4	5	6
Very bad		Bad		Good		Very good

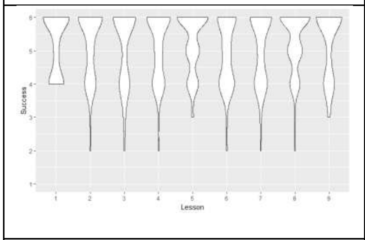
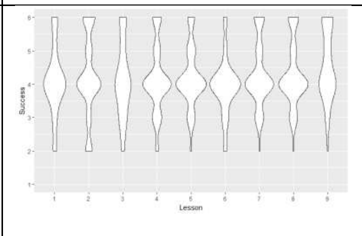
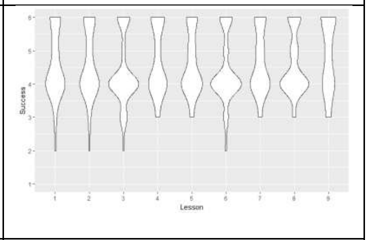
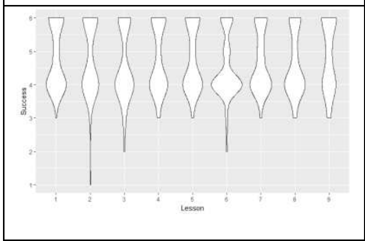
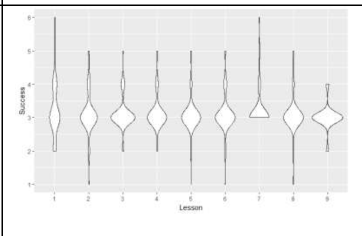
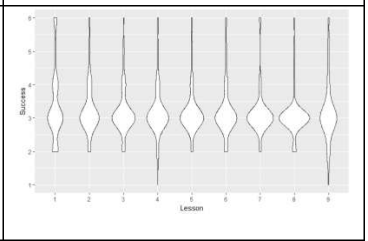
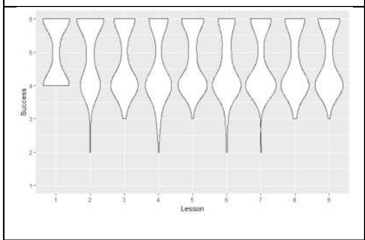
3. Rate the suitability of the *activity* for teaching and learning this lesson for your class

0	1	2	3	4	5	6
Very bad		Bad		Good		Very good

4. Rate the suitability of the *exercises* for teaching and learning this lesson for your class

0	1	2	3	4	5	6
Very bad		Bad		Good		Very good

Figure 2. Children's success in achieving lesson objectives

Appropriateness of objectives	Difficulty for children	Difficulty for teacher
		
Comfort teaching the lesson	Too much or little time preparing	Too much or little time teaching
		
Interestingness for children		
		

Nsangi A, *et al*. *BMJ Open* 2019; 9:e030787. doi: 10.1136/bmjopen-2019-030787

2. Were the objectives of this lesson appropriate for your class?

0	1	2	3	4	5	6
Inappropriate		*Somewhat inappropriate*		*Somewhat appropriate*		*Appropriate*

3. How difficult was this lesson for the children?

0	1	2	3	4	5	6
Very difficult		*Difficult*		*Easy*		*Very easy*

4. How difficult was this lesson for you?

0	1	2	3	4	5	6
Very difficult		*Difficult*		*Easy*		*Very easy*

5. How comfortable were you teaching this lesson?

0	1	2	3	4	5	6
Very uncomfortable		*Uncomfortable*		*Comfortable*		*Very comfortable*

7. Was that too little or too much time?

0	1	2	3	4	5	6
Much too little		*Too little*	*About right*	*Too much*		*Much too much*

9. Was that too little or too much time?

0	1	2	3	4	5	6
Much too little		*Too little*	*About right*	*Too much*		*Much too much*

10. How interesting was this lesson for the children?

0	1	2	3	4	5	6
Very uninteresting		*Uninteresting*		*Interesting*		*Very interesting*

Nsangi A, *et al*. *BMJ Open* 2019; 9:e030787. doi: 10.1136/bmjopen-2019-030787

Figure 3. Time used by lesson

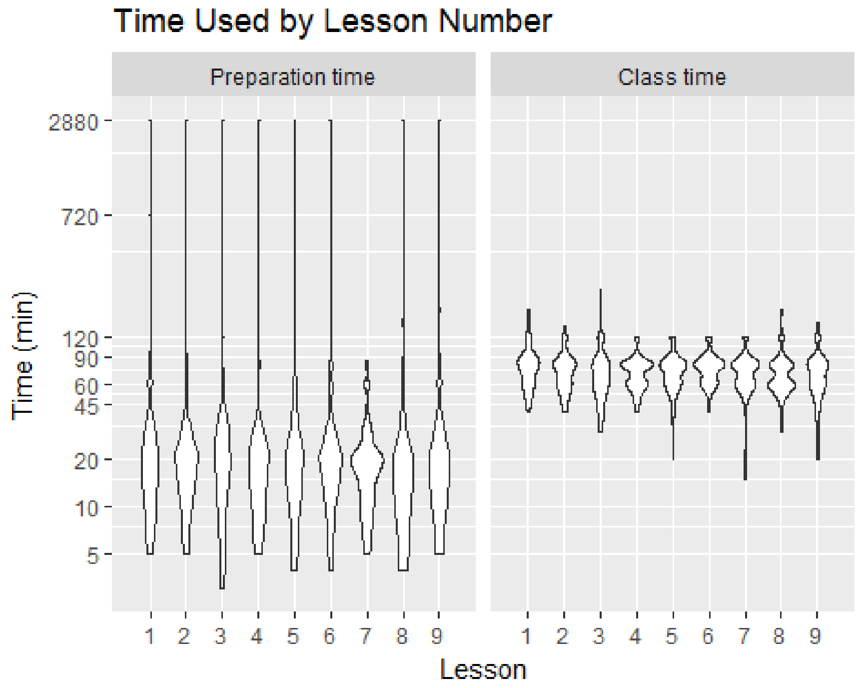
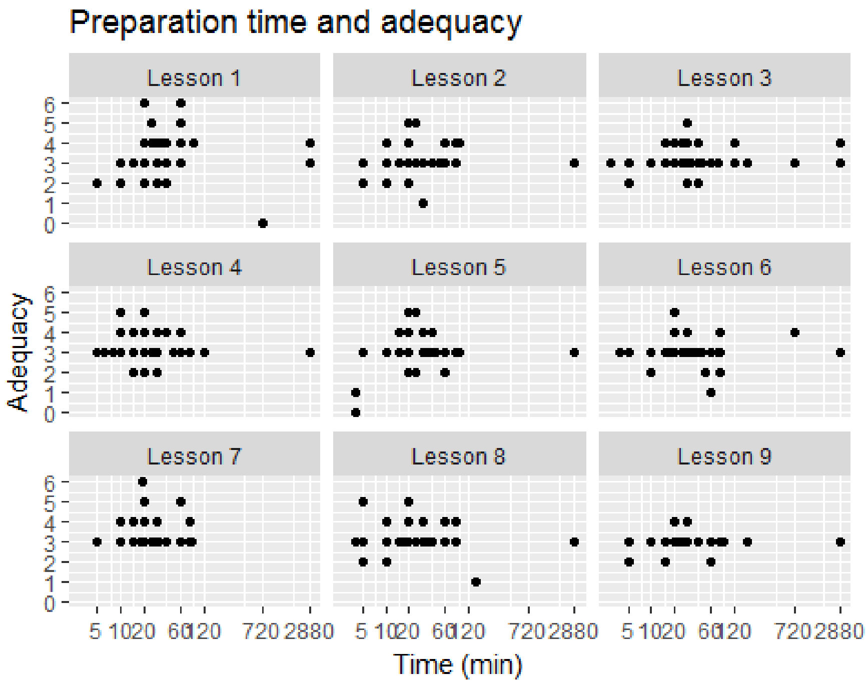
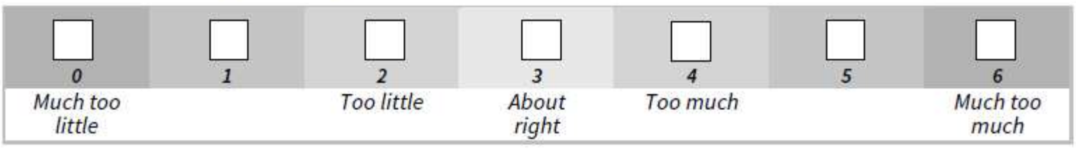


Figure 4. Preparation time in relation to teachers' assessments of whether that was too little or too much time for each lesson



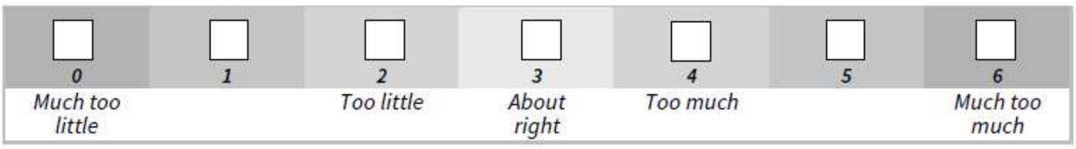
Was that too little or too much time?



Nsangi A, *et al*. *BMJ Open* 2019; 9:e030787. doi: 10.1136/bmjopen-2019-030787

Figure 5. Class time in relation to teachers' assessments of whether that was too little or too much time for each lesson

Was that too little or too much time?



Nsangi A, *et al*. *BMJ Open* 2019; 9:e030787. doi: 10.1136/bmjopen-2019-030787

Nsangi et al. Trials (2020) 21:27
https://doi.org/10.1186/s13063-019-3960-9

Trials

RESEARCH Open Access

Effects of the Informed Health Choices primary school intervention on the ability of children in Uganda to assess the reliability of claims about treatment effects, 1-year follow-up: a cluster-randomised trial

Allen Nsangi[1,2], Daniel Semakula[1,2], Andrew D. Oxman[2,3*], Astrid Austvoll-Dahlgren[3], Matt Oxman[3], Sarah Rosenbaum[3], Angela Morelli[4], Claire Glenton[3], Simon Lewin[3,5], Margaret Kaseje[6], Iain Chalmers[7], Atle Fretheim[2,3], Yunpeng Ding[3] and Nelson K. Sewankambo[1]

Abstract

Introduction: We evaluated an intervention designed to teach 10- to 12-year-old primary school children to assess claims about the effects of treatments (any action intended to maintain or improve health). We report outcomes measured 1 year after the intervention.

Methods: In this cluster-randomised trial, we included primary schools in the central region of Uganda that taught year 5 children (aged 10 to 12 years). We randomly allocated a representative sample of eligible schools to either an intervention or control group. Intervention schools received the Informed Health Choices primary school resources (textbooks, exercise books and a teachers' guide). The primary outcomes, measured at the end of the school term and again after 1 year, were the mean score on a test with two multiple-choice questions for each of the 12 concepts and the proportion of children with passing scores.

Results: We assessed 2960 schools for eligibility; 2029 were eligible, and a random sample of 170 were invited to recruitment meetings. After recruitment meetings, 120 eligible schools consented and were randomly assigned to either the intervention group (n = 60 schools; 76 teachers and 6383 children) or the control group (n = 60 schools; 67 teachers and 4430 children). After 1 year, the mean score in the multiple-choice test for the intervention schools was 68.7% compared with 53.0% for the control schools (adjusted mean difference 16.7%; 95% CI, 13.9 to 19.5; P < 0.00001). In the intervention schools, 3160 (80.1%) of 3943 children who completed the test after 1 year achieved a predetermined passing score (≥ 13 of 24 correct answers) compared with 1464 (51.5%) of 2844 children in the control schools (adjusted difference, 39.5%; 95% CI, 29.9 to 47.5).

Conclusion: Use of the learning resources led to a large improvement in the ability of children to assess claims, which was sustained for at least 1 year.

Trial registration: Pan African Clinical Trial Registry (www.pactr.org), PACTR201606001679337. Registered on 13 June 2016.

Keywords: Evidence-based healthcare, Training, Critical thinking, Health literacy, Informed decision-making, Public involvement, Children

* Correspondence: oxman@online.no
[2]University of Oslo, Oslo, Norway
[3]Centre for Informed Health Choices, Norwegian Institute of Public Health, Postboks 222 Skøyen, 0213 Oslo, Norway
Full list of author information is available at the end of the article

Summary Box

What is already known

- There is an overload of unsubstantiated claims about the benefits and harms of treatments.
- Many people are unable to assess the reliability of these claims.
- This leads to poorly informed decisions, unnecessary suffering and waste.

What are the new findings

- Children (aged 10–12 years) who used the Informed Health Choices primary school resources learned to think critically about treatment claims and retained what they learned for at least 1 year.

How might it impact clinical practice in the foreseeable future?

- In the short term, children are likely to think more critically about treatment claims and choices.
- In the long term, they may be enabled to make well-informed decisions as patients and future health professionals and as citizens and future policymakers.

Background

We identified Informed Health Choices (IHC) key concepts that people need to understand and apply when assessing claims about treatments [1, 2]. Together with teachers in Uganda, we determined which of those concepts were relevant for primary school children [3]. We then prototyped, user-tested and piloted learning resources to teach 12 key concepts (Table 1) to children [5], and we developed and validated a test to measure their ability to apply those concepts [6–10].

The resulting learning resources, which were printed in English, included a textbook, a teachers' guide, an exercise book, a poster, and cards for an activity. The textbook [11] consists of a story in a comic book format (Fig. 1), instructions for classroom activities, exercises, a checklist summarising the concepts in the book, and a glossary of keywords with definitions in English and translations to Luganda and Swahili. In addition to the textbooks, we provided intervention schools with a guide [4] for each teacher, an exercise book for each child, a poster of the checklist for the classroom, and activity cards for the seventh lesson [12]. The contents of the book and the teachers' guide are shown in Table 2. While most teachers considered the IHC content to be new, many found the design of the IHC lessons to be compatible with their teaching styles, particularly the use of multiple examples in the teachers' guide [13]. We did not intervene in the control schools.

Table 1 Twelve key concepts covered by the Informed Health Choices primary school resources

Claims
• Treatments may be harmful. • Personal experiences or anecdotes (stories) are an unreliable basis for assessing the effects of most treatments. • Widely used treatments or treatments that have been used for a long time are not necessarily beneficial or safe. • New, brand-named, or more expensive treatments may not be better than available alternatives. • Opinions of experts or authorities do not alone provide a reliable basis for deciding on the benefits and harms of treatments. • Conflicting interests may result in misleading claims about the effects of treatments.
Comparisons
• Evaluating the effects of treatments requires appropriate comparisons • Apart from the treatments being compared, the comparison groups need to be similar (i.e., 'like needs to be compared with like'). • If possible, people should not know which of the treatments being compared they are receiving. • Small studies in which few outcome events occur are usually not informative, and the results may be misleading. • The results of single comparisons of treatments can be misleading.
Choices
• Treatments usually have beneficial and harmful effects.

The concepts are shown here as they are described in the key concepts list [3], which was not designed as a learning resource, not as they were presented to the children in the primary school resources [4]

We conducted a cluster-randomised trial to evaluate the effects of using the learning resources [14, 15]. The intervention included a 2-day introductory workshop for the teachers, as well as providing them with the learning resources. The trial showed that the intervention resulted in a large improvement in the ability of children to assess claims about the effects of treatments, measured at the end of the term during which the intervention was delivered [14]. In this paper, we report on outcomes measured 1 year after the intervention. We report a process evaluation in a separate paper [13].

Methods

Details regarding the study methods can be found in the trial protocol [15] and report of the initial results [14]. They are briefly summarised here.

Participants

Between April 11, 2016, and June 8, 2016, we randomly selected 170 of 2029 eligible schools in central Uganda and recruited 120 of those schools (Fig. 2). We randomly sampled schools proportionately from lists of randomly selected districts, stratifying by school ownership (private or public) and location (urban, semi-urban and rural). We excluded international schools, special needs schools for children with visual and audio impairments, schools that had participated in user testing and piloting of the resources, infant and nursery schools and adult education schools. We included all year 5 children in the eligible schools.

Fig. 1 An excerpt from the comic story in the textbook

Random allocation and blinding

We randomly allocated schools to the intervention or control group using a computer-generated sequence. We used stratified randomisation to help ensure equal distribution of schools for school ownership (public or private) and geographical location (urban, semi-urban or rural). Research assistants labelled opaque envelopes with the unique codes, inserted cards with the study group allocated to each code in the envelopes, and sealed them. After obtaining consent from 120 schools,

Table 2 Contents of the textbook and the teachers' guide

Health Choices Book Learning to think carefully about treatments A health science book for primary school children	Teachers' Guide
Introduction • Lesson 1: Health, treatments and effects of treatments John and Julie learn about CLAIMS about treatments • Lesson 2: Someone's experience using a treatment • Lesson 3: Other bad bases for claims about treatments (part 1) • Lesson 4: Other bad bases for claims about treatments (part 2) John and Julie learn about COMPARISONS of treatments • Lesson 5: Comparisons of treatments • Lesson 6: Fair comparisons of treatments • Lesson 7: Big-enough fair comparisons of treatments John and Julie learn about CHOICES about treatments • Lesson 8: Advantages and disadvantages of a treatment Review • Lesson 9: Review of what is most important to remember from this book	The teacher's guide includes an introduction to the project and the resources, and the following for each lesson, in addition to the embedded chapter from the textbook: • The objective of the lesson • A lesson preparation plan • A lesson plan • A list of materials that the teacher and children will need • A synopsis of the story • Keywords in the chapter • Review questions to ask the children after reading the story • Extra examples for illustrating the concepts • Background about examples used in the story • Teacher instructions for the classroom activity • Answers and explanations for the activity • Answers and explanations for the exercises • Background information, examples and keyword definitions for teachers

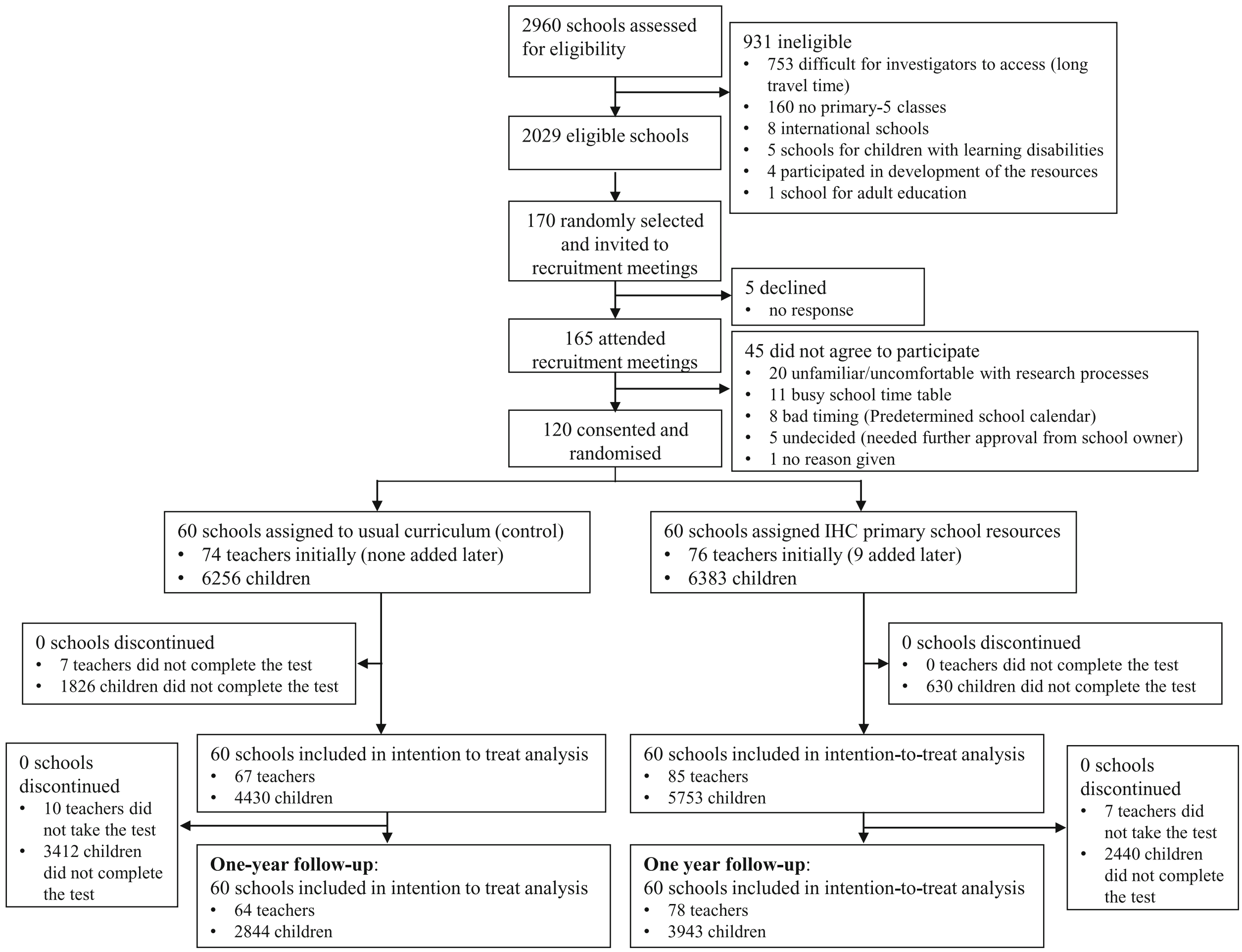


Fig. 2 Informed Health Choices trial profile

2 research assistants selected each school from a list of the schools; identified the appropriate randomisation list to be used for that school, based on its geographical location and ownership; and assigned the next available code from that list.

We informed the participating head teachers and year 5 teachers about the objectives of the study [15]. After randomisation, they knew whether they were in the intervention or control arm. The consent form stated that the outcome measure consisted of 'multiple-choice questions that assess an individual's ability to apply concepts that people must be able to understand and apply to assess treatment claims and to make informed healthcare choices.' We did not show them the test until the end of the school term. Children in both arms of the trial were informed of the purpose of the test when their teachers asked them to complete it at the end of the term and again after 1 year.

Interventions

We designed the learning resources to be used over 9 weeks, with one double-period (80 min) per week, during a single school term, and 1 h to complete the test at the end of the term and again after 1 year. There was an expectation on the part of the head teachers and teachers that any content displaced by the lessons would be compensated, so that time was not taken away from other lessons. Each school decided how to do this, with some schools using the library lessons while boarding schools preferred to teach in the evenings and on weekends [13]. The intervention was delivered between June and August 2016.

We invited all participating teachers in the intervention group to attend an introductory workshop. At the workshop, we informed them about the study objectives and procedures, including the general nature of the outcome measure; went through all nine lessons outlined in the primary school resources; and addressed any questions or concerns that arose.

We invited year 5 teachers in the control schools to a 2-h introductory meeting in each district. At these meetings, we informed them about the study procedures, including the general nature of the test that we would be using as the outcome measure. We told them that they would receive the primary school resources at the end of the study. We did not introduce them to the resources.

Outcomes

The primary outcomes, measured using the same test taken at the end of the term when the intervention was delivered, were as follows:

1. The mean test score (percentage of correct answers) on the same test 1 year later
2. The proportion of children with a passing score

Secondary outcomes were as follows:

1. The proportion of children with a score indicating mastery of the concepts
2. For each concept, the proportion of children who answered both questions correctly
3. The children's intended behaviours and self-efficacy
4. Self-reported behaviours
5. Mean scores, passing scores and mastery scores for the teachers, who took the same test as the children at the end of the intervention term and again 1 year later

Most teachers completed the test at the same time as the children. We contacted teachers who were not available on the day of the exam to arrange completion of the questionnaire on another day. The children and the teachers were aware that missing answers would be scored as wrong.

The test included 24 multiple-choice questions (2 for each concept) (Additional file 1) [9]. The questions had between two and four response options, with an overall probability of answering 39% of the questions correctly by chance alone. Two additional multiple-choice questions were included because the test used in this trial was also used in a linked randomised trial evaluating a podcast given to the parents of some of the children [16]. These two extra questions were not included in the primary analyses.

The test also included questions that assessed intended behaviours, self-efficacy, attitudes and reading skills (Additional file 1). For questions about intended behaviours and self-efficacy, we dichotomised the responses (e.g., very unlikely or unlikely versus very likely or likely) in the analysis, and we reported the proportions of children for each of the four responses. We used the answers to the reading skills questions as a covariate in exploratory analyses. In the test taken after 1 year, we also collected data on self-reported behaviours (Table 3). We made the comparisons shown in Additional file 2: Table S1 with the corresponding hypotheses. These were not specified in the original protocol for the study but were planned prior to collecting the 1-year follow-up data.

Children were counted as 'passing' or 'failing' depending on whether they met a pre-specified passing score (a minimum of 13 of 24 questions answered correctly) [6]. We used a second cut-off for a score that indicated mastery of the 12 concepts (a minimum of 20 of 24 questions answered correctly).

We also report attendance and scores on national examinations for the intervention term and for the following term. These comparisons were originally planned as part of the process evaluation [13]. We asked participating schools to provide us with school attendance records and summary score sheets containing all pupils' end-of-intervention term examination scores. The summary score sheet (Table 4) contains percentage scores for each end-of-intervention term examination and a total score across subjects (Additional file 2: Table S2). The children receive marks for English, mathematics, social studies, and science. We measured the mean difference between the intervention and control groups for each subject and for their total score (out of 100). We hypothesised higher scores in the intervention schools for English (because of the time spent reading and learning new words in English during the IHC lessons), science (based on results of randomised trials of other interventions to teach critical thinking [17–19], and stimulation of interest in science), and average scores (due to expected higher scores in English and science).

We asked teachers to record unexpected adverse events and problems that might pose risks to the children or others and to report these to the investigators or to the Institutional Review Board at Makerere University College of Health Sciences. Teachers in the intervention arm of the trial were given instructions for recording adverse events and problems in journals that they were asked to keep [13].

Statistical analysis

Statistical analysis was carried out using the University of Aberdeen Health Services Research Unit's Cluster Sample Size Calculator, applying the following assumptions: 70 children per cluster; an intraclass correlation coefficient (ICC) of 0.5, based on ICCs from a meta-analysis of randomised trials of school interventions and an international comparison of ICCs for educational achievement outcomes [20, 21]; 0% as the proportion of children expected to achieve a passing score without our intervention, based on findings from pilot testing; 10% as the smallest difference we wanted to be able to detect; an alpha of 0.05; and a power of 90%. On the basis of these assumptions, we estimated that we would need a minimum of 55 schools in each arm to detect a

Table 3 Comparisons related to self-reported behaviours in the 1-year follow-up

Question	Hypothesis and basis for the hypothesis
How often do you hear treatment claims?	Children in the intervention group will report hearing treatment claims more often because of being more aware of treatment claims and identifying them when they are made.
[For the last treatment claim that you heard,] did you think about what that treatment claim that you heard was based on?	A larger proportion of children in the intervention group will answer yes because of being more aware that many claims do not have a reliable basis.
How sure are you that the treatment claim you heard is true or can be trusted?	A smaller proportion of children in the intervention group will answer 'very sure' or 'I don't know', and a larger proportion of children in the intervention group will answer this question consistently with their answer to the preceding question about the basis of the claim (Table 5) because of being better able to assess the trustworthiness of claims and many claims not having a reliable basis.
How sure are you about the advantages and disadvantages of the [most recent] treatment you used?	A higher proportion of the children in the intervention group will answer 'not very sure because I only know about the advantages', and a smaller proportion will answer 'very sure', because information about the disadvantages of treatments is often lacking. However, this difference, if there is one, will likely be small, because children in the intervention group are more likely to consider and seek information about the disadvantages of treatments.
Who do you think should decide for you whether you should use a treatment or not use a treatment?	A higher proportion of the children in the intervention group will answer that they want to be included (A, C, D, F or G) because of having learned about how to make informed health choices; and that someone who knows a lot about treatments should be included (E, F or G), because of being more aware of the importance of assessing the reliability of evidence of effects and the skills that are needed to do this. However, this difference, if there is one, will likely be small, because children in the intervention group are more likely to recognise that expert opinion alone is not a reliable basis for a claim about treatment effects.
What happens if the claim that comes in is about negative effects of the treatment?	A larger proportion of children in the intervention group will answer, 'Not very sure because there was not a good reason behind the claims about the advantages of the treatment', because they are more likely to identify a claim whose basis was bad.
Given your thoughts about the basis of the claim, what did you yourself decide to do about the treatment?	A smaller proportion of children in the intervention group versus the control group would choose to use a treatment (in question 29.7) having recognised that the basis of the claim was untrustworthy (in question 29.6)

difference of 10% in the proportion of children with a passing score [14].

For the primary and secondary outcomes, we used mixed models with a random effects term for the clusters and the stratification variables modelled as fixed effects, using logistic regression for dichotomous outcomes and linear regression for continuous outcomes. The statistical analyses were performed with R version 3.3.2 software (R Core Team, Vienna, Austria). We used a penalized-maximum likelihood logistic regression (R package 'logistf') for the secondary outcome of passing scores for teachers because of rare events (only one teacher in the intervention group did not have a passing score). We converted odds ratios from logistic regression analyses to adjusted differences using the intervention group percentage as the reference. All the children and teachers who completed the test were included in the analyses.

Table 4 Ranges of marks and points awarded for each subject

Exam score (out of 100)	Points awarded	Marks
80–100	1	Distinction 1
70–79	2	Distinction 2
65–69	3	Credit 3
60–64	4	Credit 4
55–59	5	Credit 5
50–54	6	Credit 6
45–49	7	Pass 7
35–44	8	Pass 8
Below 35	9	Failure

For questions about intended behaviours and self-efficacy, we dichotomised the responses in the analysis and reported the proportions of children for each of the four response options. For comparisons of how frequently participants in both groups reported hearing treatment claims, we analysed the data as ordinal data using mixed ordinal logistic regression, and we dichotomised the responses.

User testing of the questions about self-reported behaviours by 40 children who did not participate in the trial suggested that the questions are understood by children in Uganda. In addition, we used open-ended questions to ensure that the children understood these questions correctly: 'Please write down the treatment claim that you last heard. What did they say the treatment would change or not change about someone's health?' (Table 5). We coded answers to these questions

Table 5 Consistent (correct) answers regarding certainty about treatment claims[a]

If you heard about a treatment claim, what was it based on?	How sure are you that the treatment claim you heard is true or can be trusted?
Someone's personal experience using the treatment	Not very sure, because the reason behind the claim was not good
What an expert said about it	Not very sure, because the reason behind the claim was not good
A research study that compared the treatment with another treatment or no treatment	Not very sure, because the reason behind the claim was not good OR Very sure, because the reason behind the claim was good
Something else	Not very sure, because the reason behind the claim was not good
I could not tell what the treatment claim was based on	Not very sure, because I don't know the reason behind the claim

[a]Questions 28.5 and 28.6 in Additional file 1

as 'correct' or 'incorrect', and we excluded from the comparisons in (Table 6) all participants who did not correctly identify the type of treatment (Additional file 2: Table S3) or who did not report a treatment claim. For attendance, we compared rates in the intervention and control groups. For marks, we compared mean exam scores (Additional file 2: Table S5), the proportions of children with passing scores (≥ 35), and the proportions of children with distinction scores (≥ 70).

To explore the risk of bias due to attrition, which was larger in the control schools than in the intervention schools, we conducted two sensitivity analyses. First, we conducted an analysis using inverse probability weighting. In this, the children in each school were given a weight equal to the inverse of the proportion of children in the school who had completed the test. Second, using the Lee bounds approach [22], we calculated upper and lower bounds for the mean difference in test scores. The bounds are constructed by trimming the group with less attrition at the upper and lower tails of the outcome (test score) distribution, respectively. In this analysis, the sample was trimmed in the intervention schools so that the proportion of children included in the analysis was equal for both groups. We did not adjust for covariates in the sensitivity analysis.

We conducted two additional sensitivity analyses to explore why the effects for the primary outcomes were smaller after 1 year than they were at the end of the intervention term. First, we calculated the differences between effects (adjusted mean differences and odds ratios) between the first and second tests based on mixed models with a random effects term for the clusters (schools) and individuals (who are used twice in these analyses), and the stratification variables modelled as fixed effects, using linear regression for the mean scores and logistic regression for the proportions with a passing score. Second, we estimated the effects at the end of the intervention term, excluding children who did not take the second test, using the same model as described above.

We explored whether there were differences in the effect of the intervention for children with advanced reading skills (all four literacy questions answered correctly) versus basic reading skills (both basic literacy questions correct and one or two of the advanced literacy questions wrong) versus lacking basic reading skills (one or both basic literacy questions wrong). In order to put the effect of the intervention in the context of effect sizes reported for other interventions to improve critical thinking or learning in primary schools [23], we calculated

Table 6 Exclusion criteria for self-reported behaviours

Response options for questions 28.2 and 29.3	Response to questions 28.3 and 29.4
28.2 What was the treatment in the claim you last heard about?	28.3 Please write the claim that you last heard.
29.3 What was the treatment for which you or an adult made the decision?	What was the claim about the treatment for which you or an adult made the decision?
Using a medicine (e.g., taking a tablet or syrup)	Exclude if the claim is not about a medicine
Getting an operation (e.g., removing a bad tooth)	Exclude if the claim is not about an operation
Using something to feel better or to heal more quickly (e.g., using a bandage or glasses)	Exclude if the claim is not about equipment
Something else (eating food or drinking something to feel better; e.g., herbs or fruit)	Exclude if the claim is not about eating/drinking something (e.g., herbs or fruit)
Avoiding doing something to feel better (e.g., not drinking milk)	Exclude if the claim is not about avoiding something
Something else	Exclude if the claim is not about a treatment ('anything done to care for yourself, so you stay well or, if you are sick or injured, so you get better and not worse')

the adjusted standardised mean difference (Hedges' *g*) and its 95% confidence interval using formulae described by White and Thomas [24].

Parents of 675 children in either the intervention or control group were recruited to participate in a parallel trial [16]. That trial evaluated the effects of a podcast designed to teach the parents of primary school children nine IHC key concepts, eight of which were included in the primary school resources. We conducted a second subgroup analysis to explore whether having a parent who listened to the podcast improved the scores of the children and whether there was an interaction between the effect of the podcast and the primary school resources. Because the parents allocated to listen to the podcast did not do so until after the children had completed the tests the first time, we only conducted this analysis for the 1-year follow-up study. We used statistical models as described above for this analysis; the main effects of the podcast were also included in these analyses.

Results

All 120 schools that were randomised provided data for the primary outcome measures and were included in the primary analyses. Most of the schools in both groups were urban or semi-urban (Table 7). There were more public schools in the control group (55% versus 50%). For the 1-year follow-up, there were fewer teachers who taught science as their main subject. Teachers in Ugandan primary schools frequently move and switch the major subject that they teach due to changes in staffing. Therefore, changes in the main subject taught by teachers are not unusual. There were otherwise only minor differences in the characteristics of the participants between the end of the intervention term and the 1-year follow-up, and between the intervention and control groups.

Six intervention schools had more than one year 5 class (with a different teacher for each class). This resulted in nine more teachers receiving training and being included in the intervention schools. No teachers were added in the control schools, because the teachers in the control schools did not receive training. For the 1-year follow-up, 78 (92%) of the teachers in the intervention schools and 59 (88%) of the teachers in the control schools completed the same test that the children took at the end of the term.

Table 7 Characteristics of the participants

		One-year follow-up		End of intervention term	
		Control schools	Intervention schools	Control schools	Intervention schools
Schools (selected from the central region of Uganda)		$N = 60$	$N = 60$	$N = 60$	$N = 60$
Location	Rural	8 (13%)	6 (10%)	8 (13%)	6 (10%)
	Semi-urban	15 (25%)	14 (23%)	15 (25%)	14 (23%)
	Urban	37 (62%)	40 (67%)	37 (62%)	40 (67%)
Ownership	Public	33 (55%)	30 (50%)	33 (55%)	30 (50%)
	Private	27 (45%)	30 (50%)	27 (45%)	30 (50%)
Teachers[a]		$N = 74$	$N = 85$	$N = 74$	$N = 85$
Completed tests		59 (80%)	78 (92%)	67 (91%)	85 (100%)
Education	Certificate	27 (46%)	34 (44%)	30 (45%)	39 (46%)
	Diploma	31 (53%)	35 (45%)	33 (49%)	35 (41%)
	University degree	1 (2%)	9 (12%)	3 (4%)	10 (12%)
Main subject taught	Science	32 (54%)	48 (56%)	49 (73%)	68 (80%)
Sex	Women	24 (41%)	32 (45%)	29 (43%)	34 (40%)
Children (enrolled in year-5 at the start of the term)		$N = 6256$	$N = 6383$	$N = 6256$	$N = 6383$
Completed tests[b]		2844 (45%)	3943 (62%)	4430 (71%)	5753 (90%)
Completed tests per school[c]	Median (25th to 75th percentile) (Range)	40 (24 to 57) (4 to 300)	49 (30 to 77) (10 to 270)	60 (40 to 95) (12 to 150)	61 (43 to 89) (18 to 176)
Sex	Girls	1558 (55%)	2164 (55%)	2457 (55%)	3154 (55%)
Age	Median (25th to 75th percentile) (Range)	12 (10 to 14) (9 to 18)	12 (10 to 14) (8 to 19)	11 (10 to 12) (8 to 20)	11 (10 to 12) (8 to 18)

[a]The number of teachers who completed the test at the end of the first term. Head teachers were initially asked to identify teacher who taught science to children in the fifth year of primary school. However, some schools had more than one year 5 class. Six intervention schools with more than one year 5 class (with a different teacher for each class) requested that nine additional teachers be included altogether
[b]Questions about the characteristics of the teachers and children were included in the test completed at the end of the school term and 1 year later
[c]The average number of year 5 children per school at the start of the term was 84 in both groups

Altogether, 6787 children completed the 1-year follow-up test (Table 7). As was the case with the test taken at the end of the intervention term, more children completed the follow-up test in the intervention schools (62%) than in the control schools (45%). We think this is because teachers who taught the lessons were more motivated to arrange for the children whom they had taught to take the test. The proportion of girls (55%) and the median age of children (12 years; 25th to 75th percentile, 10 to 14) in the two groups were the same. Most of the children answered all the questions. The proportion of missing values (unanswered questions) for each question was between 0.25% and 3.38%, and the number of missing values was similar between the intervention and control schools (Additional file 2: Table S4).

Only 64 schools provided data on the secondary outcome of school attendance. Ninety-three schools provided data on examination scores for the intervention term, and 99 provided data for the following term (Additional file 2: Table S5).

Primary outcomes and sensitivity analyses

The average score for children in the intervention schools was 68.7% compared with 53.0% in the control schools (Table 8). The adjusted mean difference (based on the regression analysis) was 16.7% (95% CI, 13.9% to 19.5%; $P < 0.00001$) higher in the intervention than in the control group. The distribution of test scores is shown in Additional file 3. In the intervention schools, 80.1% of the children had a passing score (≥ 13 of 24 correct answers) compared with 51.5% in the control schools (Table 8). The adjusted difference (based on the odds ratio from the logistic regression analysis) was 39.5% more children who passed (95% CI, 29.9% to 47.5%) in the intervention group than in the control group. Although the average score and the proportion of children with a passing score in the intervention group were higher after 1 year than at the end of the intervention term, the differences between the intervention and control schools were smaller, because the scores increased more in the control schools than in the intervention schools between the first and second tests.

We conducted two sensitivity analyses to investigate possible explanations for the small effect estimates after 1 year. To explore whether the apparent differences might have occurred by chance alone, we calculated the probability of a difference as large as or larger than what we observed having occurred by chance (Additional file 2: Table S18). It is highly unlikely that the differences in the effect estimates would have occurred by chance ($P >$ 0.00001). To explore whether the differences might reflect bias resulting from differential loss to follow-up, we calculated the effects at the end of the intervention term, excluding children who did not take the second test (Additional file 2: Table S19). The effect estimates are similar. We consider other possible explanations in the 'Discussion' section below.

We conducted two sensitivity analyses to assess the potential risk of bias from attrition (i.e., children who did not take the test) (Table 9). There was very little difference between the results of analysis using inverse probability weighting and the primary analysis (Additional file 2: Table S6), suggesting that the results are robust. In the second analysis, we calculated Lee bounds for the mean difference in test scores. This resulted in lower (worst case) and upper (best case) mean differences of 6.4% and 26.6%, respectively (95% CI, 6.6% to 26.5%). This indicates that even with the worst-case scenario, the average test score in the intervention schools was still 6.4% higher than in the control schools. Moreover, the worst-case scenario, which removed 17% of the children with the highest test scores from the intervention group, is unlikely. This is equivalent to assuming that the children in the control schools who did not take the test would have had scores that corresponded to the top 17% of the children in the intervention schools, had they taken the test (Additional file 2: Table S7). It is more likely that the children who were lost to follow-up and did not take the test would have done worse rather than better than the children who did take the test.

Secondary outcomes

- *The proportion of children with a score indicating mastery of the concepts*

In the intervention schools, 28.9% of the children had a score indicating mastery of the 12 key concepts (≥ 20 of 24 correct answers) compared with 4.9% of the children in the control schools (Table 8). The adjusted difference was 25.0% more children in the intervention schools who mastered the concepts (95% CI, 23.2% to 26.5%). This is a larger difference than there was at the end of the term during which the intervention had been delivered (18.0%). The proportion of children with a score indicating mastery increased from 18.6% to 28.9% in the intervention group between the first and second tests, compared with an increase from 0.9% to 4.9% in the control group.

- *For each concept, the proportion of children who answered both questions correctly*

For each concept, the proportion of children who answered both questions correctly was higher in the intervention schools than in the control schools, including for the concept that was not covered in the primary

Table 8 Main test score results at 1-year follow-up

	Control schools	Intervention schools	Adjusted difference[a]	Odds ratio[a]	ICC
Primary outcome					
One-year follow-up					
Mean score, %	Mean score: 53.0% (SD 16.8%)	Mean score: 68.7% (SD 18.2%)	Mean difference: 16.7% (95% CI, 13.9% to 19.5%) $P < 0.00001$		0.18
End of intervention term					
Mean score, %	Mean score: 43.1% (SD 15.2%)	Mean score: 62.4% (SD 18.8%)	Mean difference: 20.0% (95% CI, 17.3% to 22.7%)		0.18
One-year follow-up					
Passing score (≥ 13 of 24 correct answers)	51.5% of children ($n = 1464/2844$)	80.1% of children ($n = 3160/3943$)	39.5% more children (95% CI, 29.9% to 47.5%)	5.88 (95% CI, 4.00 to 8.33) $P < 0.00001$	0.20
End of intervention term					
Passing score (≥ 13 of 24 correct answers)	26.8% of children ($n = 1186/4430$)	69.0% of children ($n = 3967/5753$)	49.8% more children (95% CI, 43.8% to 54.6%)	9.34 (95% CI, 6.62 to 13.18)	0.19
Secondary outcomes					
One-year follow-up					
Mastery score (≥ 20 of 24 correct answers)	4.9% of children ($n = 139/2844$)	28.9% of children ($n = 1138/3943$)	25.0% more children (23.2–26.5%)	10.00 (95% CI, 6.67 to 16.67) $P < 0.00001$	0.19
End of intervention term					
Mastery score (≥ 20 of 24 correct answers)	0.9% of children ($n = 38/4430$)	18.6% of children ($n = 1070/5753$)	18.0% more children (95% CI, 17.5% to 18.2%)	35.33 (95% CI, 20.58 to 60.67)	0.21
Teachers' scores					
One-year follow-up					
Mean score, %	Mean score: 68.5% (SD 14.9%)	Mean score: 86.2% (SD 10.2%)	Mean difference: 17.5% (13.2% to 21.8%) $P < 0.00001$		
End of intervention term					
Mean score, %	Mean score: 66.7% (SD 14.3%)	Mean score: 84.6% (SD 17.1%)	Mean difference: 18.3% (95% CI, 12.9% to 23.3%)		
One-year follow-up					
Passing score (≥ 13 of 24 correct answers)	85.9% of teachers ($n = 50/59$)	98.7% of teachers ($n = 77/78$)	9.4% more teachers (1.3% to 52.0%)	9.12[b] (95% CI, 2.01 to 86.7) $P = 0.003$	
End of intervention term					
Passing score (≥ 13 of 24 correct answers)	86.6% of teachers ($n = 58/67$)	97.6% of teachers ($n = 83/85$)	11.3% more teachers (95% CI, 4.0% to 13.0%)	7.24 (95% CI, 1.49 to 35.26)	
One-year follow-up					
Mastery score (≥ 20 of 24 correct answers)	22.0% of teachers ($n = 13/59$)	67.9% of teachers ($n = 53/78$)	46.5% more teachers (28.1% to 61.3%)	7.70 (95% CI, 3.56 to 17.70) $P < 0.00001$	
End of intervention term					
Mastery score (≥ 20 of 24 correct answers)	14.9% of teachers ($n = 10/67$)	71.8% of teachers ($n = 61/85$)	56.7% more teachers (95% CI, 37.3% to 70.4%)	14.38 (95% CI, 6.24 to 33.14)	

[a]The adjusted difference is based on mixed models with a random effects term for the clusters (for the children only) and the stratification variables modelled as fixed effects, using logistic regression for dichotomous outcomes and linear regression for continuous outcomes. The odds ratios from the logistic regressions have been converted to differences based on the intervention school proportions and the odds ratios calculated using the intervention schools as the reference (the inverse of the odds ratios shown here)

[b]Penalized-maximum likelihood logistic regression (R package 'logistf') was used for this analysis because of rare events (only one teacher in the intervention group did not have a passing score)

Table 9 Sensitivity analyses at 1-year follow-up

	Adjusted difference[a]	Odds ratio
Mean score		
Primary analysis	Mean difference: 16.7% (95% CI, 13.9% to 19.5%) $P < 0.00001$	
Weighted analysis	Mean difference: 16.7% (95% CI, 13.9% to 19.5%)	
Lee bounds	6.4% to 26.6% (95% CI, 6.6% to 26.5%)	
Passing score (≥ 13 of 24 correct answers)		
Primary analysis	39.5% (95% CI, 29.9% to 47.5%)	5.88 (95% CI, 4.00 to 8.33) $P < 0.0001$
Weighted analysis	40.9% (95% CI, 31.0% to 49.4%)	6.25 (95% CI, 4.17 to 9.09) $P < 0.0001$

[a]The adjusted difference is based on mixed models with a random effects term for the clusters and the stratification variables modelled as fixed effects, using logistic regression for dichotomous outcomes and linear regression for continuous outcomes. The odds ratios from the logistic regressions for passing scores have been converted to differences based on the intervention school proportions and the odds ratios calculated using the intervention schools as the reference (the inverse of the odds ratios shown here)

school resources ($P < 0.0001$ for all 13 concepts after a Bonferroni correction for multiple comparisons) (Table 10).

- *Children's intended behaviours and self-efficacy*

Compared with children in the control schools, children in the intervention schools were more likely to respond that they would find out the basis for a claim (adjusted difference, 8.1%; 95% CI, 3.7% to 12.6%) and to participate in a research study if asked (adjusted difference, 7.7%; 95% CI, 2.0% to 13.5%) (Additional file 2: Table S8). These findings are similar to those we found 1 year earlier. However, there was little if any difference in how likely they were to find out if a claim was based on research (adjusted difference, 2.6%; 95% CI, – 1.9% to 7.2%). This contrasts with what we found 1 year earlier (10.8%; 95% CI, 6.3% to 15.1%).

- *Self-reported behaviours*

Similar to what we found 1 year earlier, children in the intervention schools were more likely to consider it easy to assess whether a claim is based on research than children in the control schools (adjusted difference, 14.8%; 95% CI, 8.9% to 20.5%) (Table 11). They were also more likely to consider it easy to find information about treatments based on research (adjusted difference, 7.2%; 95% CI, 2.6% to 11.5%) (Table 12), whereas 1 year earlier, we had detected little if any difference (Additional file 2: Table S9). We detected little if any difference in how easy children thought it was to assess how sure they could be about the results of research or to assess how relevant research findings are to them. One year earlier, compared with children in the control group, the children in the intervention group were less likely to consider it easy to assess how sure they could be about the results of research.

The children in the intervention schools were more likely to report hearing one or more treatment claims daily or weekly (Table 13) than were children in the control schools (adjusted difference, 7.0%; 95% CI, 0.5% to 12.9%) (Additional file 2: Table S10). The children in the intervention schools were less likely to be very sure or not to know whether a claim could be trusted (Table 14) (adjusted difference, – 15%; 95% CI, – 9.9% to – 19.7%) and more likely to assess the trustworthiness of a claim consistently with what they identified as the basis of the claim (adjusted difference, 7.6%; 95% CI, 3.5% to 11.1%) (Additional file 2: Table S11). However, there were only slight differences in how likely children in the intervention schools were to think about the basis of the last claim that they heard (Table 15) (adjusted difference, 4.1%; 95% CI, – 1.2% to 9.6%) (Additional file 2: Table S12 and S13), as well as in their assessments of the advantages and disadvantages of the most recent treatment they had used (Table 16) (Additional file 2: Table S14). The difference in attendance or examination scores was also small (Additional file 2: Table S5). As reported previously [14], none of the teachers or research assistants who observed the lessons reported any adverse events.

- *Mean, passing and mastery scores for teachers*

After 1 year, most teachers in both the intervention and control groups (98.7% and 85.9%, respectively) had passing scores (adjusted difference, 8.6%; 95% CI, 1% to 55.5%) (Table 8). The teachers in the intervention group were much more likely to have a score indicating mastery of the concepts (67.9% versus 21.9%; adjusted difference, 46.3%; 95% CI, 31.5% to 56.6%). These results are similar to those we found at the end of the intervention term.

Subgroup analyses

As was the case at the end of the intervention term, the intervention still had positive effects 1 year later, regardless of reading skills (Table 17), but with larger effects for children with better reading skills (Additional file 2: Table S15). Compared with the control schools (Table 18), reading skills were better in the intervention schools at the end of the intervention term and after 1 year (Additional file 2: Table S16). They had improved by about the same amount in both the intervention and control schools after 1 year. We did not detect an interaction between having a parent who listened to the podcast and the primary school intervention (Table 19) (adjusted difference for the interaction,

Table 10 Results for each concept for children at 1-year follow-up

No.	Concept	Control schools % correct[a] No. of schools = 60 No. of children = 2844	Intervention schools % correct[a] No. of schools = 60 No. of children = 3943	Adjusted difference[b] (95% CI)	ICC[c]	Odds ratio (95% CI)
	Claims					
1.1	Treatments may be harmful.	40.5% (n = 1152)	64.6% (n = 2547)	29.2% (22.4–35.0%)	0.120	3.33 (2.50–4.35) P < 0.00001
1.2	Personal experiences or anecdotes (stories) are an unreliable basis for assessing the effects of most treatments.	26.5% (n = 753)	52.0% (n = 2052)	30.0% (24.5–34.2%)	0.119	3.85 (2.86–5.00) P < 0.00001
1.3	A treatment outcome may be associated with a treatment, but not caused by the treatment.[d]	27.3% (n = 776)	36.4% (n = 1436)	11.2% (6.4–15.2%)	0.087	1.69 (1.33–2.13) P = 0.00002
1.4	Widely used treatments or treatments that have been used for a long time are not necessarily beneficial or safe.	26,3% (n = 748)	54,4% (n = 2144)	30.0% (23.8–35.1%)	0,157	3.70 (2.70–5.00) P < 0.00001
1.5	New, brand-named, or more expensive treatments may not be better than available alternatives.	48.9% (n = 1392)	73.6% (n = 2901)	28.1% (22.2–34.5%)	0.088	3.33 (2.63–4.35) P < 0.00001
1.6	Opinions of experts or authorities do not alone provide a reliable basis for deciding on the benefits and harms of treatments.	43.2% (n = 1230)	67.6% (n = 2664)	26.8% (20.3–33.3%)	0.113	3.03 (2.33–4.00) P < 0.00001
1.7	Conflicting interests may result in misleading claims about the effects of treatments.	37.0% (n = 1051)	47.2% (n = 1861)	10.8% (5.5–15.9%)	0.077	1.56 (1.25–1.96) 0.00009
	Comparisons					
2.1	Evaluating the effects of treatments requires appropriate comparisons.	10.3% (n = 294)	32.0% (n = 1263)	24.2% (21.1–26.2%)	0.148	5.56 (3.85–7.69) P < 0.00001
2.2	A part from the treatments being compared, the comparison groups need to be similar (i.e., 'like needs to be compared with like').	12.1% (n = 344)	29.3% (n = 1155)	16.6% (14.2–18.9%)	0.063	2.86 (2.33–3.57) P < 0.00001
2.5	If possible, people should not know which of the treatments being compared they are receiving.	23.3% (n = 664)	36.2% (n = 1428)	15.1% (11.4–18.8%)	0.070	2.13 (1.72–2.70) P < 0.00001
3.1	Small studies in which few outcome events occur are usually not informative and the results may be misleading.	32.6% (n = 928)	50.3% (n = 1984)	20.5% (15.8–25.3%)	0.082	2.38 (1.92–3.03) P < 0.00001
4.1	The results of single comparisons of treatments can be misleading.	29.1% (n = 827)	44.8% (n = 1766)	17.6% (12.4–22.2%)	0.096	2.17 (1.69–2.78) P < 0.00001
	Choices					
5.1	Treatments usually have beneficial and harmful effects.	35.2% (n = 1000)	50.8% (n = 2004)	16.8% (11.4–22.1%)	0.090	2.00 (1.59–2.56) P < 0.00001

[a]There were two multiple-choice questions for each concept. The proportions are for the percentage of children who answered both questions correctly
[b]The adjusted difference is based on mixed models with a random effects term for the clusters and the stratification variables modelled as fixed effects, using logistic regression. The odds ratios from the logistic regressions have been converted to differences based on the intervention school proportions and the inverse of the odds ratios shown here
[c]Intraclass correlation coefficient
[d]This concept was not included in the learning resources or counted in the average, pass or mastery scores

3.8%; 95% CI, − 3.9% to 11.4%) (Additional file 2: Table S17).

Discussion

The large effect that the Informed Health Choices intervention had on the ability of primary school children in Uganda to assess claims about treatment effects was sustained after 1 year. The mean score and the proportions of children with passing and mastery scores increased in the intervention schools (Table 8). However, because the scores in the control schools increased more than the scores in the intervention schools, the differences between

Table 11 Intended behaviours at 1-year follow-up

Think about an illness that you might get. Imagine someone claiming (saying) that a particular treatment might help you get better.

	How likely are you to find out what the claim was based on (e.g., by asking the person making the claim)?		How likely are you to find out if the claim was based on a research study comparing the treatment with no treatment (a fair comparison)?		How likely are you to say 'yes' if you are asked to participate in a research study comparing two treatments for your illness (a fair comparison)?	
	Control schools $N = 2844$	Intervention schools $N = 3943$	Control schools $N = 2844$	Intervention schools $N = 3943$	Control schools $N = 2844$	Intervention schools $N = 3943$
Missing	69 (2.4%)	67 (1.7%)	87 (3.1%)	70 (1.8%)	36 (1.3%)	44 (1.1%)
Very unlikely	217 (7.6%)	376 (9.5%)	301 (10.6%)	467 (11.8%)	245 (8.6%)	277 (7.0%)
Unlikely	289 (10.2%)	376 (9.5%)	424 (14.9%)	569 (14.4%)	329 (11.6%)	429 (10.9%)
Likely	975 (34.3%)	1510 (38.3%)	747 (26.3%)	997 (25.3%)	1045 (36.7%)	1577 (40.0%)
Very likely	678 (23.8%)	1082 (27.4%)	705 (24.8%)	1164 (29.5%)	719 (25.3%)	1155 (29.3%)
I don't know	616 (21.7%)	532 (13.5%)	580 (20.4%)	676 (17.1%)	470 (16.5%)	461 (11.7%)
Likely or very likely[a]	1653 (58.1%)	2592 (65.7%)	1452 (51.1%)	2161 (54.8%)	1764 (62.0%)	2732 (69.3%)
Odds ratio (95% CI)[b]	1.41 (1.18–1.69) $P = 0.00020$		1.11 (0.93–1.33) $P = 0.269$		1.41 (1.10–1.79) $P = 0.00629$	
Adjusted difference[b]	8.1% (3.7–12.6%)		2.6% (−1.9% to 7.2%)		7.7% (2.0–13.5%)	
End of intervention term[c]						
Likely or very likely	2440 (55.1%)	3731 (64.9%)	1967 (44.4%)	3114 (54.1%)	2163 (48.8%)	3201 (55.6%)
Odds ratio	1.56 (95% CI, 1.29 to 1.88)		1.54 (95% CI, 1.29 to 1.84)		1.37 (95% CI, 1.16 to 1.62)	
Adjusted difference	10.6% (95% CI, 6.2% to 14.7%)		10.8% (95% CI, 6.3% to 15.1%)		7.8% (95% CI, 3.7% to 11.9%)	

[a] Missing values and don't know are pooled with unlikely and very unlikely

[b] The difference is an adjusted difference, based on mixed models with a random effects term for the clusters and the stratification variables modelled as fixed effects, using logistic regression. The odds ratios from the logistic regressions have been converted to differences using the intervention schools as the reference and the inverse of the odds ratios shown here

[c] Results based on responses at the end of the term when the intervention was delivered

Table 12 Self-efficacy

How difficult or easy would you find each of these actions to be?

	Assessing whether a claim about a treatment is based on a research study comparing treatments (a fair comparison)		Assessing where I can find information about treatments that is based on research studies comparing treatments (fair comparisons)		Assessing how sure I can be about the results of a research study comparing treatments (the trustworthiness of the results)		Assessing if the results of a research study comparing treatments are likely to be relevant to me	
	Control schools $N = 2844$	Intervention schools $N = 3943$	Control schools $N = 2844$	Intervention schools $N = 3943$	Control schools $N = 2844$	Intervention schools $N = 3943$	Control schools $N = 2844$	Intervention schools $N = 3943$
Missing	71 (2.5%)	55 (1.4%)	73 (2.6%)	71 (1.8%)	82 (2.9%)	84 (2.1%)	72 (2.5%)	86 (2.2%)
Very difficult	357 (12.6%)	455 (11.5%)	338 (11.9%)	431 (10.9%)	488 (17.2%)	581 (14.7%)	436 (15.3%)	568 (14.4%)
Difficult	779 (27.4%)	865 (21.9%)	634 (22.3%)	876 (22.2%)	653 (23.0%)	1007 (25.5%)	513 (18.0%)	727 (18.4%)
Easy	837 (29.4%)	1517 (38.5%)	899 (31.6%)	1348 (34.2%)	640 (22.5%)	897 (22.7%)	694 (24.4%)	1027 (26.0%)
Very easy	334 (11.7%)	623 (15.8%)	525 (18.5%)	856 (21.7%)	454 (16.0%)	712 (18.1%)	562 (19.8%)	779 (19.8%)
I don't know	466 (16.4%)	428 (10.9%)	375 (13.2%)	361 (9.2%)	527 (18.5%)	662 (16.8%)	567 (19.9%)	756 (19.2%)
Easy or svery easy[a]	1171 (41.2%)	2140 (54.3%)	1424 (50.1%)	2204 (55.9%)	1094 (38.5%)	1609 (40.8%)	1256 (44.2%)	1806 (45.8%)
Odds ratio (95% CI)[b]	1.82 (1.43–2.33) $P < 0.00001$		1.33 (1.11–1.59) $P = 0.00171$		1.10 (0.94–1.30) $P = 0.233$		1.10 (0.93–1.28) $P = 0.279$	
Adjusted difference[b]	14.8% (8.9–20.5%)		7.2% (2.6–11.5%)		2.3% (– 1.4% to 6.1%)		2.3% (– 1.9% to 6.1%)	
End of intervention term[c]								
Easy or very easy	1886 (42.6%)	3244 (56.4%)	3069 (53.3%)	2238 (50.5%)	1777 (40.1%)	2112 (36.7%)	2002 (45.2%)	2727 (47.4%)
Odds ratio	1.83 (95% CI, 1.55 to 2.16)		1.13 (95% CI, 0.96 to 1.33)		0.84 (95% CI, 0.73 to 0.96)		1.08 (95% CI, 0.93 to 1.25)	
Adjusted difference	15.0% (95% CI, 10.9% to 19.0%)		3.0% (95% CI, – 1.0% to 7.0%)		– 4.1% (95% CI, – 1.0% to – 7.3%)		1.9% (95% CI, – 1.8% to 5.6%)	

[a]Missing values and don't know are pooled with difficult and very difficult

[b]The difference is an adjusted difference, based on mixed models with a random effects term for the clusters and the stratification variables modelled as fixed effects, using logistic regression. The odds ratios from the logistic regressions have been converted to differences using the intervention schools as the reference and the inverse of the odds ratios shown here

[c]Results based on responses at the end of the term when the intervention was delivered

Table 13 Self-reported behaviour: awareness of treatment claims

How often do you hear treatment claims?		
	Control schools $N = 2844$	Intervention schools $N = 3943$
One or more most days	572 (20.1%)	1000 (25.4%)
One or more most weeks	374 (13.2%)	599 (15.2%)
One or more most months	497 (17.5%)	715 (18.1%)
Almost never	653 (23.0%)	788 (20.0%)
I don't know	717 (25.2%)	810 (20.5%)
Missing	31 (1.1%)	31 (0.8%)
One or more most days or most weeks	946 (33.8%)	1599 (40.6%)
Odds ratio[a]	1.35 (95% CI, 1.02–1.79) $P = 0.0356$	
Adjusted difference[b]	7.0% (95% CI, 0.5–12.9%)	

[a]The odds ratio for the dichotomised data is shown in the table. The odds ratio from the mixed ordinal logistic regression was 1.30 (95% CI, 1.01 to 1.67; $P = 0.0431$)
[b]The difference is an adjusted difference, based on a mixed model with a random effects term for the clusters and the stratification variables modelled as fixed effects, using logistic regression. The odds ratio from the logistic regression has been converted to a difference using the intervention schools as the reference and the inverse of the odds ratios shown here

the intervention and control schools for the mean score and the proportion of children with a passing score were smaller, albeit still large. On the other hand, the difference in the proportion of children with a mastery score increased.

We considered five possible explanations for these findings, none of which seem likely. First, the apparent differences in the effect estimates between the first and second measurements is unlikely to have occurred by chance alone (Additional file 2: Table S18). Second, bias resulting from differential loss to follow-up is also unlikely to explain the differences (Additional file 2: Table S19). A third possible explanation is that there was a learning effect from taking the test the first time, which was greater in the control schools than in the intervention schools. It is possible that the learning effect of taking the test alone would be greater than the added learning effect of taking the test after having been exposed to the IHC lessons. 'Testing effects'—gains in learning that occur when students take a practice test—are well documented [25, 26]. They occur with and without feedback [26] and for higher-level thinking ('application' in Bloom's taxonomy) as well as for recall of basic facts [25]. However, most studies investigating testing effects have been conducted over a much shorter time frame [26], and we are not aware of any studies that have documented a difference in testing effects between students who studied before taking a practice test and others who did not study. A fourth possible explanation is

Table 14 Self-reported behaviour: assessment of trustworthiness of treatment claims

How sure are you that the treatment claim you heard is true or can be trusted?		
	Control schools $N = 2844$	Intervention schools $N = 3943$
Missing	49 (1.7%)	60 (1.5%)
Not very sure because I don't know the reason behind the claim	665 (23.4%)	1039 (26.4%)
Not very sure because the reason behind the claim was not good	543 (19.1%)	1087 (27.6%)
Very sure because the reason behind the claim was good	704 (24.8%)	790 (20.0%)
I don't know because I don't know how to decide whether it is true or not	883 (31.0%)	967 (24.5%)
Very sure or I don't know	1587 (55.8%)	1757 (44.6%)
Odds ratio (very sure or I don't know vs other)	0.55 (95% CI, 0.45–0.67) $P < 0.0001$	
Adjusted difference[a]	−15.0% (95% CI, − 9.9% to − 19.7%)	
Odds ratio (consistent with what they identified as the basis for the claim)[b]	1.45 (95% CI, 1.18–1.75) $P = 0.000549$	
Adjusted difference[a]	7.6% (95% CI 3.5% - 11.1%)	

[a]The differences are adjusted differences, based on mixed models with a random effects term for the clusters and the stratification variables modelled as fixed effects, using logistic regression. The odds ratio from the logistic regression has been converted to a difference using the intervention schools as the reference and the inverse of the odds ratios shown here
[b]See Table 5

Table 15 Self-reported behaviour: assessment of the basis of treatment claims

For the last treatment claim that you heard, did you think about what that treatment claim that you heard was based on?	Control schools $N = 2844$	Intervention schools $N = 3943$
Missing	50 (1.8%)	57 (1.4%)
No	512 (18.0%)	845 (21.4%)
Yes	1387 (48.8%)	2116 (53.7%)
I don't remember	895 (31.5%)	925 (23.5%)
Odds ratio (yes versus other)	1.18 (95% CI, 0.95–1.47) $P = 0.130$	
Adjusted difference[a]	4.1% (95% CI, −1.2% to 9.6%)	

[a]The difference is an adjusted difference, based on a mixed model with a random effects term for the clusters and the stratification variables modelled as fixed effects, using logistic regression. The odds ratio from the logistic regression has been converted to a difference using the intervention schools as the reference and the inverse of the odds ratios shown here

that children learn to think critically about treatment claims naturally as they grow older or through the existing curriculum, and the control schools were catching up with the intervention schools because of this. However, as documented in our process evaluation, the content of the lessons was new for all of the teachers and not something that they had previously taught. Furthermore, we did not deliver the learning resources to the control schools until after the follow-up data had been collected. Fifth, it also seems unlikely that the improvement was due to an improvement in reading skills in the control schools, because the change in reading skills was similar in the intervention and control schools.

The effects that we found for the children for each IHC key concept, as well as the effects that we found for the teachers, were similar to those we found at the end of the intervention term. Overall, these findings support the conclusion that the effects of the intervention were sustained, even though we are unable to explain why the children's scores increased more in the control schools than in the intervention schools.

Other findings provide modest support for the conclusion that the children in the intervention schools were more likely to use what they had learned. The children in the intervention schools remained more likely than those in control schools to find out the basis for a treatment claim, more confident in their ability to assess whether a treatment claim is based on research, and more likely to participate in a research study if asked. They also appeared to be somewhat more aware of treatment claims, more sceptical of treatment claims, and more likely to assess the trustworthiness of treatment claims. However, all of these differences were smaller than the difference for the primary outcome measures. Moreover, at the end of the intervention term, children in the intervention schools were more likely than children in the control schools to say they would find out if a treatment claim was based on research, but after 1 year there was little difference.

Table 16 Self-reported behaviour: assessment of advantages and disadvantages of treatments

How sure are you about the advantages and disadvantages of the [most recent] treatment you used?	Control schools $N = 2844$	Intervention schools $N = 3943$
A. Not very sure because I don't know the reasons behind the claims about the good and bad things that treatment makes happen	531 (18.7%)	851 (21.6%)
B. Not very sure because there was not a good reason behind the claims about the advantages of the treatment	355 (12.5%)	549 (13.9%)
C. Not very sure because I only know about the advantages of the treatment. I also need to know about the disadvantages	765 (26.9%)	992 (25.2%)
D. Very sure because there is a good reason behind the claims about the advantages and disadvantages of the treatment	652 (22.9%)	929 (23.6%)
E. I did not use any treatment	498 (17.5%)	590 (15.0%)
Missing	43 (1.5%)	32 (0.8%)
Odds ratio (C versus any other response)	1.05 (95% CI, 0.86–1.30) $P = 0.62$	
Adjusted difference answer C vs else	−0.9% (95% CI, −5.3% to 2.7%)	
Odds ratio (D versus any other response)	1.03 (95% CI, 0.85–1.23) $P = 0.79$	
Adjusted difference answer D vs else	−0.5% (95% CI, −3.9% to 2.8%)	

Table 17 Subgroup analysis: reading skills[a]

	Control schools	Intervention schools	Adjusted difference[b]	Odds ratio	ICC
Mean score, %					
Lacking basic reading skills ($N = 1775$)	No. of children = 893	No. of children = 882			
	Mean score: 47.2% (SD 16.4%)	Mean score: 57.1% (SD 18.1%)	Mean difference: 11.2% (95% CI, 8.2% to 14.2%)		0.146
Basic reading skills ($N = 2672$)	No. of children = 1093	No. of children = 1579			
	Mean score: 55.2% (SD 16.9%)	Mean score: 67.9% (SD 16.8%)	Mean difference: 14.8% (95% CI, 12.3% to 17.3%)		0.162
Advanced reading skills ($N = 2340$)	No. of children = 858	No. of children = 1482			
	Mean score: 56.3% (SD 15.6%)	Mean score: 76.5% (SD 15.5%)	Mean difference: 19.4% (95% CI, 16.9% to 21.9%)		0.117
Passing score (≥ 13 of 24 correct answers)					
Lacking basic reading skills ($N = 1775$)	No. of children = 893	No. of children = 882			
	36.6% of children $n = 327$	59.3% of children $n = 523$	28.9% more children (95% CI, 20.8% to 36.7%)	0.30 (95% CI, 0.20 to 0.43)	0.144
Basic reading skills ($N = 2672$)	No. of children = 1093	No. of children = 1579			
	57.0% of children $n = 623$	81.2% of children $n = 1282$	33.6% more children (95% CI, 24.0% to 41.9%)	0.21 (95% CI, 0.15 to 0.31)	0.150
Advanced reading skills ($N = 2340$)	No. of children = 858	No. of children = 1482			
	60.0% of children $n = 514$	91.4% of children $n = 1355$	33.4% more children (95% CI, 25.7% to 42.5%)	0.13 (95% CI, 0.09 to 0.18)	0.098
Mastery score (≥ 20 of 24 correct answers)					
Lacking basic reading skills ($N = 1775$)	No. of children = 893	No. of children = 882		0.22	
	3.0% of children $n = 27$	10,1% of children $n = 89$	7.7% more children (95% CI, 5.6% to 8.8%)	(95% CI, 0.12 to 0.42)	0.220
Basic reading skills ($n = 2672$)	No. of children = 1093	No. of children = 1579		0.15	
	6.5% of children $n = 71$	24.1% of children $n = 380$	19.6% more children (95% CI, 17.0% to 21.3%)	(95% CI, 0.09 to 0.24)	0.192
Advanced reading skills ($n = 2340$)	No. of children = 858	No. of children = 1482		0.06	
	4.8% of children $n = 41$	45.1% of children $n = 669$	40.4% more children (95% CI, 38.2% to 41.9%)	(95% CI, 0.04 to 0.09)	0.139

[a]Because reading skills were measured after the intervention, we have not reported a test of interaction here (see Additional file 2)
[b]The adjusted difference is based on mixed models with a random effects term for the clusters and the stratification variables modelled as fixed effects, using logistic regression for dichotomous outcomes and linear regression for continuous outcomes. The odds ratios from the logistic regressions for passing scores and mastery scores have been converted to differences using the intervention school proportions and the inverse of the odds ratios shown here

Table 18 Differences in reading skills

Reading skills	Immediately after the intervention[a]			One-year follow-up[a]			Change from first to second test[a]		
	Control schools No. of children 4412 *n* (%)	Intervention schools No. of children 5711 *n* (%)	Diff	Control schools No. of children 2844 *n* (%)	Intervention schools No. of children 3943 *n* (%)	Diff	Control schools	Intervention schools	Diff
Lacking basic reading skills	2139 (48.5%)	2224 (38.9%)	−9.5%	893 (31.4%)	882 (22.4%)	−9.0%	−17.1%	−16.6%	0.5%
Basic reading skills	1507 (34.2%)	2155 37.7%	3.6%	1093 (38.4%)	1579 (40.0%)	1.6%	4.3%	2.3%	−2.0%
Advanced reading skills	766 (17.4%)	1332 23.3%	6.0%	858 (30.2%)	1482 (37.6%)	7.4%	12.8%	14.3%	1.5%

[a]Reading skills as measured by first four questions in the test administered at the end of the term when the intervention was delivered and the same test 1 year later. The differences (Diff) are shown between the intervention and control schools for each time the test was administered and the change from the first to the second time

The data we were able to collect for attendance and national examinations were incomplete, but based on those data, there was little difference between children in the intervention and control schools (Table 20). This contrasts with findings of studies in the United Kingdom, which have shown beneficial effects of critical thinking or meta-cognition interventions on academic achievement [17–19]. Possible explanations for this include the limitations of the data we were able to collect for these outcomes and differences between the interventions and the contexts in which they were delivered.

The main limitations of our follow-up study are similar to those discussed in our report of effects found immediately after the intervention [14]. First, we cannot rule out some degree of bias due to attrition. However, sensitivity analyses suggest that the effect estimates are robust. Second, we used an outcome measure that we developed ourselves. Outcome measures developed by the study authors for use in a study may be more likely to find larger effects than studies using established measures of critical thinking [23]. We developed the outcome measure because there was no pre-existing outcome measure suitable for our study [8]. Although we have demonstrated the validity and reliability of the outcome measure [6, 7, 9, 10], one should be cautious about comparing our results with the effects of other critical thinking interventions. Moreover, we are unaware of any other directly comparable studies [20, 23, 27–30]. Other interventions in primary schools have been found to improve critical thinking [23], but these studies have been conducted in high-income countries, few have measured outcomes after 1 year, and neither the interventions nor the outcome measures are directly comparable [27, 29].

It remains uncertain how transferable the findings of this study are to other countries. However, pilot testing in Kenya, Norway and Rwanda suggest that it may be possible to use the IHC primary school resources without substantial modifications. They have already been translated to Kiswahili, Kinyarwanda, Spanish, French and Farsi. There are plans or expressions of interest to translate them to other languages, including Chinese, German and Italian. Pilot studies have been completed or planned in several other countries, including Ireland and South Africa. The resources are open access, and we have prepared a guide for translating, contextualising and testing them [31].

However, we believe that a one-off intervention is unlikely to have large long-term effects on decision-making, health behaviours or health. Rather, we view this as the first step in developing a set of interventions for a spiral curriculum [32, 33]. Using this approach, some of the IHC key concepts would be introduced, as we did in this study. Then those concepts would be reinforced in subsequent cycles, and other, more complex concepts would be introduced.

Table 19 Subgroup analysis: parent who listened to the podcast

	Control schools	Intervention schools	Adjusted effect of the interaction[a]
Parent in control group ($N = 167$)	No. of children = 69 Mean score: 55.1% (SD 16.4%)	No. of children = 98 Mean score: 64.5% (SD 20.2%)	Mean difference: 3.8% (95% CI, − 3.9% to 11.4%) $P = 0.3443$
Parent in podcast group ($N = 168$)	No. of children = 64 Mean score: 53.6% (SD 15.9%)	No. of children = 104 Mean score: 66.3% (SD 18.6%)	

[a]Adjusted for location, ownership (public/private) and random effect of clustering, ICC = 0.185

Table 20 Attendance and national examinations

Attendance rates				
	Control schools *N* = 33 schools Median (25th to 75th percentile)	Intervention schools *N* = 31 schools Median (25th to 75th percentile)	Adjusted difference	*P* value
Intervention term	90.3% (78.7% to 98.0%)	89.1% (80.4% to 96.4%)	3% less (95% CI, −14 to 6)	0.437
Following term	91.7% (81.1% to 97.8%)	89.5% (78.6% to 96.2%)	2% more (95% CI, −10 to 13)	0.726
Average scores on national examinations				
	Control schools Mean (SD)	Intervention schools Mean (SD)	Adjusted mean difference	*P* value
End of intervention term				
English	54.2% (22.5)	52.3% (22.5)	−1.7% (95% CI, −6.6 to 3.2)	0.500
Math	51.5% (23.4)	49.0% (22.5)	−1.8% (95% CI, −6.6 to 3.0)	0.457
Science	49.8% (24.4)	49.7% (23.3)	−0.5% (95% CI, −5.4 to 4.5)	0.852
Social science	52.6% (24.0)	51.9% (23.7)	−1.0% (95% CI, −6.2 to 4.2)	0.699
Total	52.3% (21.4)	51.1% (21.0)	−1.2% (95% CI, − 5.5 to 3.2)	0.597
Following term				
English	56.3% (22.1)	56.1% (22.5)	2.4% (95% CI, −2.3 to 7.2)	0.312
Math	53.8% (23.2)	50.2% (22.4)	0.8% (95% CI, −4.1 to 5.8)	0.752
Science	52.4% (23.9)	49.3% (23.3)	0.8% (95% CI, − 4.1 to 5.4)	0.813
Social science	56.0% (23.8)	52.0% (22.7)	−0.1% (95% CI, −4.8 to 4.7)	0.964
Total	54.8% (21.5)	52.2% (20.6)	1.0% (95% CI, −3.4, 5.4)	0.671
Proportion with a passing score (≥ 35%) on the national examinations				
	Control schools *n* (%)	Intervention schools *n* (%)	Adjusted difference	
End of intervention term	Total: 49 schools, 3795 children	Total: 44 schools, 4201 children		
English	2917/3766 (77.5%)	3009/3984 (71.8%)	0.0% (95% CI, −10.0 to 13.8)	0.998
Math	2709/3772 (71.8%)	2809/3985 (70.5%)	1.6% (95% CI, −12.0 to 11.9)	0.799
Science	2632/3764 (69.9%)	2829/3990 (70.9%)	−0.1% (95% CI, −11.4 to 14.6)	0.988
Social science	2794/3773 (74.1%)	2957/3980 (74.3%)	−1.7% (95% CI, − 11.9 to 12.9)	0.801
Total	2698/3730 (72.3%)	2830/3934 (71.9%)	−0.7% (95% CI, − 11.5 to 13.8)	0.920
Following term	Total: 51 schools, 3956 children	Total: 48 schools, 4474 children		
English	3205/3934 (81.5%)	3655/4460 (82.0%)	3.8% (95% CI, −5.2 to 16.6)	0.461
Math	3038/3940 (76.9%)	3174/4441 (71.5%)	−0.1% (95% CI, −10.3 to 12.8)	0.984
Science	2923/3942 (74.2%)	3137/4436 (70.7%)	−0.1% (95% CI, −11.4 to 14.6)	0.878
Social science	3125/3940 (79.3%)	3366/4452 (75.6%)	1.1 (95% CI, −8.1 to 13.2)	0.839
Total	3022/3914 (77.2%)	3268/4404 (74.2%)	1.5% (95% CI, −8.6 to 14.8)	0.797
Proportion with a distinction score (≥ 70%) on the national examinations				
	Control schools *n* (%)	Intervention schools *n* (%)	Adjusted difference	
End of intervention term	Total: 49 schools, 3795 children	Total: 44 schools, 4201 children		
English	1133/3766 (30.1%)	1077/3984 (27.0%)	−7.0% (95% CI, −21.4 to 4.9)	0.278
Math	995/3772 (26.4%)	850/3985 (21.3%)	−4.2% (95% CI, −17.3 to 5.6)	0.716
Science	966/3764 (25.7%)	977/3990 (24.5%)	−2.1% (95% CI, −14.9 to 7.7)	0.716
Social science	1117/3773 (29.6%)	1117/3980 (28.1%)	−1.7% (95% CI, − 15.5 to 9.2)	0.791
Total	904/3730 (24.2%)	882/3934 (22.4%)	−2.1% (95% CI, − 15.0 to 7.3)	0.693
Following term	Total: 51 schools, 3956 children	Total: 48 schools, 4474 children		
English	1263/3934 (32.1%)	1440/4460 (32.3%)	4.8% (95% CI, −7.7 to 14.6)	0.425
Math	1101/3940 (27.9%)	1023/4441 (23.0%)	−3.4% (95% CI, −16.8 to 6.6)	0.551

Table 20 Attendance and national examinations *(Continued)*

Attendance rates				
	Control schools *N* = 33 schools Median (25th to 75th percentile)	Intervention schools *N* = 31 schools Median (25th to 75th percentile)	Adjusted difference	*P* value
Science	1099/3942 (27.9%)	1024/4436 (23.1%)	−0.8% (95% CI, −12.3 to 7.9)	0.875
Social science	1342/3940 (34.1%)	1207/4452 (27.1%)	−0.2% (95% CI, − 12.4 to 9.3)	0.967
Total	1063 (27.2%)	1012 (23.0%)	1.3% (95% CI, −11.1 to 10.0)	0.819

SD standard deviation

Conclusions

It is possible to teach young children in a low-income country to think critically about the trustworthiness of claims about the benefits and harms of treatments, and children retain what they have learned for at least 1 year. In this study, we were also able to document modest effects on self-reported behaviours, because young children seldom make actual health choices independently. We believe it is highly desirable to begin teaching the IHC key concepts at a young age, and we have shown that this is possible.

Supplementary information

Supplementary information accompanies this paper at https://doi.org/10.1186/s13063-019-3960-9.

Additional file 1. The claim evaluation tools.

Additional file 2: Table S1. Comparisons related to self-reported behaviours in the 1-year follow-up. **Table S2.** Ranges of marks and points awarded for each subject. **Table S3.** Exclusion criteria for self-reported behaviours. **Table S4.** Number of missing values for each question. **Table S5.** Attendance and national examinations. **Table S6.** Sensitivity analyses – 1-year follow-up. **Table S7.** Attrition, differences in test scores across strata of schools. **Table S8.** Intended behaviours – 1-year follow-up. **Table S9.** Self-efficacy. **Table S10.** Self-reported behaviour – awareness of treatment claims. **Table S11.** Self-reported behaviour – assessment of trustworthiness of treatment claims. **Table S12.** Consistent (correct) answers regarding certainty about treatment claims. **Table S13.** Self-reported behaviour – assessment of the basis of treatment claims. **Table S14.** Self-reported behaviour – assessment of advantages and disadvantages of treatments. **Table S15.** Subgroup analysis – reading skills. **Table S16.** Differences in reading skills. **Table S17.** Subgroup analysis – parent who listened to the podcast. **Table S18.** Exploratory analyses – *P* values for differences between first (end of intervention term) and second (1-year follow-up) effects. **Table S19.** Exploratory analyses excluding children who did not take the test both times.

Additional file 3. Distribution of scores and curves.

Abbreviations

IHC: Informed Health Choices Project

Acknowledgements

We are grateful to the Global Health and Vaccination Research (GLOBVAC) programme of the Research Council of Norway for support of this research and to the UK National Institute for Health Research for supporting Iain Chalmers and the James Lind Initiative. This work was also partially supported by a Career Development Award from the DELTAS (Developing Excellence in Leadership, Training and Science) Africa Initiative (grant DEL-15-011 to THRiVE-2). The DELTAS Africa Initiative is an independent funding scheme of the African Academy of Sciences (AAS)'s Alliance for Accelerating Excellence in Science in Africa (AESA) and supported by the New Partnership for Africa's Development Planning and Coordinating Agency (NEPAD Agency) with funding from the Wellcome Trust (grant 107742/Z/15/Z) and the UK government. The views expressed in this publication are those of the author(s) and not necessarily those of AAS, NEPAD Agency, Wellcome Trust or the UK government. Alun Davies, Lena Nordheim, Peter O. Okebukola, Newton Opiyo, Jonathan Sharples, Helen Wilson and Charles Shey Wiysonge determined the cut-off scores for passing and mastery. Miriam Grønli was responsible for the textbook colouring, and Nora Rosenbaum assisted. Aisha Hashi, Sara Jaber, Rida Shah and Katie Tveiten helped test prototypes. Michael Mugisha, Anne-Marie Uwitonze and Jenny Moberg helped with piloting and user-testing an earlier version of the learning resources. We thank Dr. Daniel Nkaada at the Ugandan Ministry of Education for technical guidance; Sarah Natunga at the National Curriculum Development Centre in Uganda for reviewing the materials; Martin Mutyaba, Esther Nakyejwe, Margaret Nabatanzi, Hilda Mwebaza, Peter Lukwata, Rita Tukahirwa, David Ssimbwa, Adonia Sebulime, Enock Steven Ddamulira and Solomon Segawa for their help with data management; and all the research assistants who helped with data collection and entry. We also thank the Informed Health Choices advisory group. We are especially grateful to the many teachers and children in Uganda, Kenya, Rwanda and Norway who helped with the development of the Informed Health Choices primary school resources and to all the children, teachers and head teachers at the schools who participated in this trial.

Authors' contributions

AN and DS are the principal investigators. They drafted the protocol with help from the other investigators and were responsible for the day-to-day management of the trial. NKS and ADO had primary responsibility for overseeing the trial. MO and SR had primary responsibility for developing the primary school resources. AM shared primary responsibility for developing the teachers' guide. All the investigators other than KYD contributed to the development of the resources and to the protocol. AAD had primary responsibility for developing and validating the outcome measure. AN and DS had primary responsibility for data collection. KYD did the statistical analysis. The Norwegian Institute of Public Health, recipient of the grant from the Research Council of Norway, is the coordinating centre for the Informed Health Choices project. ADO, SR, AAD and IC are principal members of the coordinating group for the trial and, together with NKS and the principal investigators, acted as the steering committee for the trial. They were responsible for final decisions about the protocol and reporting of the results. All the investigators including CG, SL, MK and AF reviewed the manuscript, provided input, and agreed on the final version for publication. YD did the statistical analyses. All authors read and approved the final manuscript.

Funding

This trial was funded by the Research Council of Norway (project number 220603/H10). The funder had no role in the study design, data collection, data analysis, data interpretation or writing of the report. The principal investigator had full access to all the data in the study and had final responsibility for the decision to submit the manuscript for publication.

Availability of data and materials

The data files for the 1-year follow-up are available from the Norwegian Centre for Research Data (http://www.nsd.uib.no/nsd/english/index.html).

Ethics approval and consent to participate

Ethics approval was obtained from the institutional review board at Makerere University College of Health Sciences School of Medicine (reference number 2013-105) and the Uganda National Council for Science and Technology (reference number SS3328) at the beginning of the study, and renewal of approval was sought for the follow-up study. Informed consent for all grade 5 classes to participate in the trial was obtained from school heads (the head teacher or school director) and grade 5 teachers. We provided the head teacher of each school with information about the study and obtained written consent from them on behalf of their school to participate in the first trial (at the end of the intervention term) and the second trial (1-year follow up). In addition, we obtained written consent from the primary 5 (year 5 of primary school) teachers identified by the head teachers. Informed consent was not required from the children or their parents. We did not obtain assent from individual primary 5 children or consent from their parents, because the intervention posed minimal risk and no more risk than other teaching materials [34], almost none of which have been evaluated [20, 30]. Informed consent by individual children or their parents, in effect, would be meaningless once the decision to participate was taken by the head teacher and the teachers, who have the responsibility and authority to make decisions about lesson plans and the administration of tests [35]. Individual children and their parents had the same right to refuse participation as they do for any other lesson or test in primary schools.

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

Author details
[1]College of Health Sciences, Makerere University, Kampala, Uganda. [2]University of Oslo, Oslo, Norway. [3]Centre for Informed Health Choices, Norwegian Institute of Public Health, Postboks 222 Skøyen, 0213 Oslo, Norway. [4]Infodesignlab, Oslo, Norway. [5]Health Systems Research Unit, South African Medical Research Council, Cape Town, South Africa. [6]Tropical Institute of Community Health & Development, Kisumu, Kenya. [7]James Lind Initiative, Oxford, UK.

Received: 2 August 2019 Accepted: 4 December 2019
Published online: 06 January 2020

References
1. Chalmers I, Oxman AD, Austvoll-Dahlgren A, Ryan-Vig S, Pannell S, Sewankambo N, et al. Key Concepts for Informed Health Choices: a framework for helping people learn how to assess treatment claims and make informed choices. BMJ Evid Based Med. 2018;23(1):29–33.
2. Austvoll-Dahlgren A, Oxman AD, Chalmers I, Nsangi A, Glenton C, Lewin S, et al. Key concepts that people need to understand to assess claims about treatment effects. J Evid Based Med. 2015;8:112–25.
3. Nsangi A, Semakula D, Oxman DA, Sewankambo KN. Teaching children in low income countries to assess claims about treatment effects: a prioritisation of key concepts. J Evid Based Med. 2015;8(4):173–80.
4. Informed Health Choices Group. Teachers' guide for the Health Choices Book: learning to think carefully about treatments. A health science book for primary school children. Oslo: Norwegian Institute of Public Health; 2016. https://www.informedhealthchoices.org/wp-content/uploads/2016/08/IHC-V3-Teachers-guide_with-cover_Nov2016_lowres.pdf. Accessed 11 Feb 2019.
5. Nsangi A, Semakula D, Rosenbaum S, Oxman AD, Oxman M, Morelli A, et al. Development of the Informed Health Choices resources in four countries to teach primary school children to assess claims about treatment effects. 2019. In press. https://www.informedhealthchoices.org/wp-content/uploads/2016/08/IHC-primary-school-resources-development_IHC-Working-paper-2017.pdf
6. Davies A, Gerrity M, Nordheim L, Peter O, Opiyo N, Sharples J, et al. Measuring ability to assess claims about treatment effects: establishment of a standard for passing and mastery. In: IHC Working Paper. Oslo: Norwegian Institute of Public Health; 2017. http://www.informedhealthchoices.org/wp-content/uploads/2016/08/Claim-cut-off-IHC-Working-Paper-2017-01-09.pdf. Accessed 11 Feb 2019.
7. Semakula D, Nsangi A, Øystein G, Oxman AD, Sewankambo NK, A. A-D. Measuring ability to assess claims about treatment effects in English and Luganda: evaluation of multiple-choice questions from the "Claim Evaluation Tools" database using Rasch modelling. In: Informed Health Choices Working Paper, 2017. Oslo: Norwegian Institute of Public Health; 2017. https://www.informedhealthchoices.org/wp-content/uploads/2016/08/Claim-2nd-Rasch-analysis-in-Uganda-2017-03-17.pdf. Accessed 11 Feb 2019.
8. Austvoll-Dahlgren A, Nsangi A, Semakula D. Interventions and assessment tools addressing key concepts people need to know to appraise claims about treatment effects: a systematic mapping review. Syst Rev. 2016;5:215.
9. Austvoll-Dahlgren A, Semakula D, Nsangi A, Oxman AD, Chalmers I, Rosenbaum S, et al. Measuring ability to assess claims about treatment effects: the development of the 'Claim Evaluation Tools'. BMJ Open. 2017;7(5):e013184.
10. Austvoll-Dahlgren A, Guttersrud Ø, Semakula D, Nsangi A, Oxman AD, et al. Measuring ability to assess claims about treatment effects: a latent trait analysis of items from the 'Claim Evaluation Tools' database using Rasch modelling. BMJ Open. 2017;7:e013185.
11. Informed Health Choices Group. The Health Choices Book: Learning to think carefully about treatments. A health science book for primary school children. Oslo: Norwegian Institute of Public Health; 2016.
12. Primary School Resources. Informed Health Choices. https://www.informedhealthchoices.org/primary-school-resources/. Accessed 11 Feb 2019.
13. Nsangi A, Semakula D, Glenton C, Lewin S, Oxman AD, Oxman M, et al. Informed Health Choices intervention to teach primary school children in low-income countries to assess claims about treatment effects: process evaluation. BMJ Open 2019;9(9):e030787.
14. Nsangi A, Semakula D, Oxman AD, Austvoll-Dahlgren A, Oxman M, Rosenbaum S, et al. Effects of the Informed Health Choices primary school intervention on the ability of children in Uganda to assess the reliability of claims about treatment effects: a cluster-randomised controlled trial. Lancet. 2017;390(10092):374–88.
15. Nsangi A, Semakula D, Oxman AD, Oxman M, Rosenbaum S, Austvoll-Dahlgren A, et al. Does the use of the Informed Healthcare Choices (IHC) primary school resources improve the ability of grade-5 children in Uganda to assess the trustworthiness of claims about the effects of treatments: protocol for a cluster-randomised trial. Trials 2017;18:223.
16. Semakula D, Nsangi A, Oxman AD, Oxman M, Austvoll-Dahlgren A, Rosenbaum SE, et al. Effects of the Informed Health Choices podcast on the ability of parents of primary school children in Uganda to assess claims about treatment effects: a randomised controlled trial. Lancet. 2017;390(10092):389–98.
17. Education Endowment Foundation. Meta-cognition and self-regulation. The Sutton Trust Education Endowment Foundation Teaching and Learning Toolkit. London: Education Endowment Foundation; 2018. https://educationendowmentfoundation.org.uk/evidence-summaries/teaching-learning-toolkit/meta-cognition-and-self-regulation/. Accessed 11 Feb 2019
18. Hanley P, Slavin R, Elliot L. Thinking, doing, talking science: evaluation report and executive summary. London: Education Endowment Foundation; 2015. https://v1.educationendowmentfoundation.org.uk/uploads/pdf/Oxford_Science.pdf. Accessed 11 Feb 2019
19. Gorard S, Siddiqui N, See BH. Philosophy for children: evaluation report and executive summary. London: Education Endowment Foundation; 2015. https://educationendowmentfoundation.org.uk/public/files/Support/Campaigns/Evaluation_Reports/EEF_Project_Report_PhilosophyForChildren.pdf. Accessed 11 Feb 2019
20. McEwan PJ. Improving learning in primary schools of developing countries: a meta analysis of randomised experiments. Rev Educ Res. 2015;85:385–94.
21. Zopluoglu C. A cross-national comparison of intra-class correlation coefficient in educational achievement outcomes. J Meas Eval Educ Psychol. 2012;3:242–78.
22. Lee DS. Training, wages, and sample selection:estimating sharp bounds on treatment effects. Rev Econ Stud. 2009;76:1071–102.
23. Abrami PC, Bernard RM, Borokhovski E, Waddinton DI, Wade CA, Persson T. Strategies for teaching students to think critically: a meta analysis. Rev Educ Res. 2015;85:275–314.
24. White IR, Thomas J. Standardised mean differences in individually-randomised and cluster randomised trials, with applications to meta-analysis. Clin Trials. 2005;2:141–51.
25. Marsh EJ, Roediger HL, Bjork RA, Bjork EL. The memorial consequences of multiple-choice testing. Psychon Bull Rev. 2007;14:194–9.
26. Adesope OO, Trevisan DA, Sundararajan N. Rethinking the use of tests: a meta-analysis of practice testing. Rev Educ Res. 2017;(87):659–701.
27. Nordheim L, Gundersen MW, Espehaug B, Guttersrud O, Flottorp S. Effects of school based educational interventions for enhancing adolescents'

abilities in critical appraisal of health claims: a systematic review. PLoS One. 2016;11:0161485.
28. Castle JC, Chalmers I, Atkinson P, et al. Establishing a library of resources to help people understand key concepts in assessing treatment claims—the "Critical thinking and Appraisal Resource Library" (CARL). PLoS One. 2017;12:e0178666.
29. Cusack L, Del Mar CB, Chalmers I, Gibson E, Hoffmann TC. Education interventions to improve people's understanding of key concepts in assessing the effects of health interventions: a systematic review. Syst Rev. 2018;7:68.
30. Snilstveit B, Stevenson J, Phillips D, et al. Interventions for improving learning outcomes and access to education in low- and middle-income countries: a systematic review. London: International Initiative for Impact Evaluation; 2015. http://www.3ieimpact.org/en/evidence/systematic-reviews/details/259/. Accessed 11 Feb 2019
31. Informed Health Choices Group. Guide for piloting the Informed Health Choices (IHC) school resources. In: Informed Health Choices Working Paper. Oslo: Norwegian Institute of Public Health; 2017. http://www.informedhealthchoices.org/wp-content/uploads/2016/08/GUIDE-for-piloting-IHC-primary-school-resources.pdf. Accessed 11 Feb 2019.
32. Harden RM, Stamper N. What is a spiral curriculum? Med Teach. 1999;21:141–3.
33. Bruner JS. The process of education. Cambridge: Harvard University Press; 1960.
34. Kopelman LM. Minimal risk as an international ethical standard in research. J Med Philos. 2004;29:351–78.
35. Weijer C, Grimshaw JM, Taljaard M, et al. Ethical issues posed by cluster randomized trials in health research. Trials. 2011;12:100.

Publisher's Note

The Claim Evaluation Tools

This questionnaire includes multiple-choice questions about treatment claims. Please answer all questions to the best of your ability.

The questionnaire includes some words that may be unfamiliar to you:

A **TREATMENT** is anything done to care for yourself, so you stay well or, if you are sick or injured, so you get better and not worse. For example, wearing glasses (to see better).
IN LUGANDA: OBUJJANJABI

A **TREATMENT CLAIM** is something someone says about whether a treatment causes something to happen or to change. A claim can be true or can be false. For example, that wearing glasses makes you see better.
IN LUGANDA: EKINTU EKYOGERWAYOGERWA KUBY'OBUJJANJABI

A **RESEARCH STUDY** is a way to answer a question by carefully collecting information. For example, a study might be done to answer the question: Does wearing glasses make people see better?
IN LUGANDA: OKUNOONYEREZA OKWEKINNASAYANSI

RESULTS of a study are what the study found. For example, whether people who wear glasses could see better.
IN LUGANDA: EKIVAAMU MUKUNOONYEREZA

Part 1. Questions about you

1.1 How old are you? __________

1.2 Are you a:

☐ Girl

☐ Boy

1.3 At which school did you complete your P.5?

____

____

____

Part 2. Questions about claims

Instructions: Read the passage on every question then answer the question below the passage using one of the provided answers. <u>For each question, choose what you think is the best answer and write the letter for that answer in the box provided.</u>

Example

A teacher says that the children in his school run faster than the children going to school in another village.

Question: **How can the teacher be sure about this?**

Options:

A) He should ask a teacher at the other school

B) He should arrange for a running contest between the two schools

C) He should ask the children in his school what they think

D) He should ask the children in the other school what they think

Answer: B

Start

2. A doctor did a research study to find out if drinking tea keeps people from getting sick. He tossed a coin to decide who should get the tea and who should not. People who got tea went to the doctor's office every day to drink their tea. At the end of the study, people who got the tea were less likely to be sick than those who got no tea.

Based on the text above, please answer the following questions:

2.1 Who went to the doctor's office every day?

Options:

A) People who did not get tea

B) People who got tea

C) Everyone

D) People who got sick

Answer: ☐

2.2 How did the doctor decide who should get tea?

Options:

A) By tossing a coin

B) By asking people what they would like

C) The doctor gave tea to those who were more likely to be sick

D) The doctor asked people who came to his office

Answer: ☐

3. A doctor did a research study to find out if drinking tea keeps people from getting sick. He tossed a coin to decide who should get the tea and who should not. People who got tea went to the doctor's office every day to drink their tea. At the end of the study, people who got the tea were less likely to be sick than those who got no tea.

Based on the text above, please answer the following questions:

3.1 What was the treatment?

Options:

A) Tea

B) Sleep

C) The study

D) The doctor

Answer: ☐

3.2 What was the result of the study?

Options:

A) Drinking tea can help people from getting sick

B) Doctors should toss coins when doing studies

C) People should go to the doctor if they are sick

D) Not drinking tea can help people from getting sick

Answer: ☐

4. Annette sees an advert on TV for a new soap which the makers say protects people from getting skin rashes. Annette thinks that this soap must be better than other soaps for protecting her skin.

Question: **Is Annette right?**

Options:

A) No, the soap may be newer, but that does not mean that it is better than other soaps

B) Yes, the new soap is probably better than most other soaps because it is newer

C) Yes, the new soap is probably better than most other soaps because a well-known company makes it

Answer: ☐

5. Regina has an illness that makes it difficult to breathe. She hears on the radio about a medicine that has helped many people for their breathing problems.

Question: **How sure can Regina be that the medicine does not have any harms?**

Options:

A) It is not possible to say. However, medicines are rarely harmful

B) Not very sure, because all medicines may harm people as well as help them

C) Very sure, since the medicine has helped many people, it is unlikely that it also harms people

Answer: ☐

6. John has a skin rash on his leg. A shop sells several creams to treat skin rashes. John chooses a cream from a well-known company, even though it is more expensive than the other creams. John thinks the cream is more likely to heal his rash than the other creams because it is more expensive.

Question: **Is John right?**

Options:

A) No, just because the cream is expensive does not mean that it will work better than other creams

B) It is not possible to say. However, expensive creams are likely to be better because the companies spend more time making them

C) No, the cream is probably not as good as the other creams. Well-known companies are usually better at advertising

D) Yes, the company is well-known for a reason, so it is more likely to be better than creams sold by lesser-known companies

Answer: ☐

7. Two companies make two different medicines for treating stomach pain. Each of them says that their medicine is the better one.

Question: **How can you know which of the two medicines is better for stomach pain?**

Options:

A) It is not possible to say. The companies may just say their medicine is best because they want to make money

B) I would rely on the best known company; it is more likely to have the best medicine

C) I cannot trust either of the companies. They are probably both wrong

Answer: ☐

8. Dr. Kato and Dr. Semakula disagree about which medicine for stomach pain is best. Dr. Kato says his opinion is right because he has worked as a doctor for a longer time than Dr. Semakula.

Question: **Is Dr. Kato right?**

Options:

A) Yes, because Dr. Kato has worked for a long time, he has more experience than Dr. Semakula

B) Yes, because Dr. Kato has worked for a long time, he must be basing what he says on studies comparing the medicines

C) No, just because Dr. Kato has worked as a doctor for a longer time does not mean that he is basing what he says on studies that compare medicines for stomach pain

Answer: ☐

9. Habibah has pain in her ear, and she asks her brother Hassan what to do about it. He says that once, when he had a pain like that, he rinsed his ear with hot water. The next day, his ear pain was gone. Based on his experience, he says rinsing with hot water is helpful for ear pain.

Question: **Do you agree with Hassan?**

Options:

A) Yes. Because this is Hassan's experience, it is likely to be true

B) No, Hassan's experience is not enough to be sure

C) Yes, Hassan rinsed his ear with hot water and the next day his ear pain was gone

Answer:

10. Sarah has an illness. There is a medicine for it, but she is unsure if she should try it. A research study comparing the medicine with no medicine found that the medicine was helpful but also that it could be harmful. Three of Sarah's friends are giving her advice about what to do.

Question: **Which advice given to her by her friends is the best advice?**

Options:

A) She should only take the medicine if many people have tried the medicine before

B) She should only take the medicine if she thinks it will help her more than it will harm her

C) If Sarah has enough money to buy the medicine, it could not hurt to try it

Answer:

11. Dr. Acheng is an expert on treating headaches. A news reporter interviews Dr. Acheng about a new medicine. Dr. Acheng says that, in her personal experience, the new medicine is good for treating headaches.

Question: **How sure can we be that Dr. Acheng right?**

Options:

A) It is not possible to say. It depends on how long Dr. Acheng has been an expert on treating headaches

B) Not very sure. Even though Dr. Acheng is an expert, the new medicine still needs to be compared in studies with other treatments

C) Very sure. Dr. Acheng is an expert, so she knows if the new medicine is good or not based on her experience

D) Very sure. Dr. Acheng would not be interviewed by a news reporter if her advice was not good

Answer: ☐

12. Edith has a stomach pain. Edith's mother says that fruit juice is a good treatment for stomach pain. She learnt about this treatment from Edith's grandmother. Over many years, other families she knows have also used fruit juice to treat stomach pain.

Question: **Based on this, how sure can we be that fruit juice is a good treatment for stomach pain?**

Options:

A) Not very sure. Even though people have used fruit juice over many years, that does not mean that it helps stomach pain

B) Very sure. If it has worked for Edith's mother and other people who have tried it, it will probably work for her too

C) Not very sure. Edith should ask more families if they use fruit juice to treat stomach pain

Answer:

13. At David's school, some students have poor parents. The students with poor parents drink less fruit juice than the children of other parents. The students with poor parents are also more often sick. Based on this link, David thinks that people who drink fruit juice, are less likely to get sick.

Question: **Is David correct?**

Options:

A) It is not possible to say, it depends on whether or not Peter has poor parents

B) Yes, students with poor parents do not drink fruit juice and are more often sick

C) Yes, the juice is the only possible reason why the students with the poor parents are more often sick

D) It is not possible to say. There could be other reasons why students with poor parents are more often sick

Answer:

14. In a research study done by John, four people were told to do exercises every day for a month, and four people were told to eat bananas every day. At the end of the month, the people who ate bananas had more strength than those who did exercises. Based on his study, John advises his friend Mildred to eat bananas.

Question: **Mildred says that we cannot be sure about the results of John's study. Why?**

Options:

A) Because the study included so few people, the differences in strength could have happened by chance, and not because of the bananas

B) Because John should have included fewer people in his study so that he could have followed them more closely

C) Because four people is not enough, all people taking part in the study should have been told to eat the bananas

Answer: ☐

15. A new fruit drink is said to make people feel strong. Fred wanted to know if this is true, and decided to do a research study comparing people who got the new fruit drink and people who drank just water.

People in the study knew if they got the new drink or water, and Fred told them that the new fruit drink was likely to make people stronger. At the end of the study, Fred was right and those who drank the new fruit drink said they felt stronger.

Question: **Why can't we be sure about the results of Fred's study?**

Options:

A) Because all people taking part in the study should have been given the new fruit drink

B) Because people knew if they got the new fruit drink, and knowing this may have influenced how they felt

C) Because Fred should have told both groups that they could expect to feel stronger

Answer:

16. Harriet is worried about getting sick. She hears about a new research study on the radio that compared a new medicine to an old medicine. Fewer people who took the old medicine got sick compared to the people who took the new medicine.

Question: **How sure can Harriet be that the old medicine is better than the new medicine?**

Options:

A) Less sure, because Harriet needs to know the results of all other studies comparing the new medicine with the old medicine

B) More sure, because she heard about the study on the radio

C) Less sure, unless she finds another study with the same results

D) More sure, because this is a new study

Answer: ☐

17. Doctors studied people with stomach pain before and after they took a new medicine. After taking the new medicine, many people felt less pain.

Question: **Can we be sure that the new medicine is good for treating stomach pain?**

Options:

A) No, taking the new medicine should have been compared either with not taking the medicine, or with taking an older medicine

B) Yes, people were asked how much pain they felt before and after they took the new medicine

C) Yes, the study was done by doctors

Answer: ☐

18. In a research study, doctors compared two treatments for knee pain, a new and an old treatment. People were able to choose which treatment they got. Most young people chose the new treatment. At the end of the study, people who chose the new treatment had less pain.

Question: **How sure can you be that the new treatment is better for treating pain than the old treatment?**

Options:

A) Less sure, because people taking the new and old treatment were not similar

B) Less sure, because all people taking part in the study should have got both treatments

C) Less sure, because older people did not like the new treatment

Answer: ☐

19. Judith wants smoother skin. The younger girls in her school have smoother skin than the older girls. Judith thinks this is because the younger girls use cream on their skin to make the skin smoother.

Question: **Based on this link between using cream and smooth skin, is Judith correct?**

Options:

A) It is not possible to say. It depends on how many younger and older girls there are

B) It is not possible to say. There might be other differences between the younger and older girls

C) Yes, because the younger girls use cream on their skin and they have smoother skin

D) No, Judith should try using the cream herself to see if it works for her

Answer: ☐

20. Dr. Wasswa has done a research study giving a new medicine to people who were vomiting. Some of the people stopped vomiting after they got the new medicine. Dr. Wasswa says that this means that the medicine works.

Question: **Is Dr. Wasswa right?**

Options:

A) No. The people who used the medicine were not compared with similar people who did not use the medicine

B) Yes, some of the people stopped vomiting

C) No, since not all of the people stopped vomiting

Answer: ☐

Instructions: Read the passage at the top of the box. Then read the text in each row and choose what you think is the best answer by making a tick ✓ in one of the two boxes. There should be only one tick in each row.

21. When you are sick, sometimes people say that something - a treatment - is good for you. Below you will find different things people say about such treatments. **Do you agree or disagree with each of the following things being said?**		
For each thing being said below, use a tick ✓ to mark whether you "agree" or "disagree".		
Things being said:	**Agree**	**Disagree**
21.1 Peter says that if a treatment works for one person, the treatment will help others too	☐	☐
21.2 Alice says that if some people try the treatment and feel better, this means that the treatment helps	☐	☐
21.3 Habibah says that, just because many people are using the treatment, this does not mean that it helps	☐	☐
21.4 Julie says that companies sometimes say that the treatment they make is best just to make money	☐	☐

22. A doctor wanted to know if a new medicine for treating headaches is better than an older medicine. The doctor did a research study, comparing the two medicines. **Which of the actions would help us be more sure about the results?**
For each action below, use a tick ✓ to mark whether you think the action would help us be "more sure" or "less sure" about the results of the study.

Actions:	**More sure**	**Less sure**
22.1 The doctor should use chance (like tossing a coin) to decide which people should be given the new and which the old medicine	☐	☐
22.2 People should not know which medicine they get (the new medicine or the old medicine) until the end of the study	☐	☐
22.3 The doctor should include only a small number of people in the study	☐	☐

23. To know if a treatment helps you, the treatment should be compared in research studies to other treatments (fair comparisons). Below you will find different things people say about such studies.

Do you agree or disagree with each of the following things being said?

For each thing being said below, use a tick ✓ to mark whether you "agree" or "disagree".

Things being said:	**Agree**	**Disagree**
23.1 Julie says that, if a treatment has been compared in a study to another treatment, you don't have to look for more studies	☐	☐
23.2 Margaret says that the results of a study should be used to decide if a treatment is more helpful than harmful	☐	☐

Part 3. Questions about your views

Instructions: For the following questions, there are no right or wrong answers. Read the passage at the top of the box. Then read the text in each row and <u>choose what you think is the best answer by making a tick ✓ in one of the five boxes</u>. There should be only one tick in each row.

24. Think about an illness that you might get. Imagine someone claiming (saying) that a particular treatment might help you get better.

How likely are you to do each of the following actions?

(Mark with a tick ✓ in the box; one check for each row.)

Actions:	**Very unlikely**	**Unlikely**	**Likely**	**Very likely**	**I don't know**
24.1 Find out what the claim was based on (for example by asking the person making the claim)	☐	☐	☐	☐	☐
24.2 Find out if the claim was based on a research study comparing the treatment to no treatment (a fair comparison)	☐	☐	☐	☐	☐

25. Below are some actions. Please read each one carefully and give the answer that comes closest to **how difficult or easy you find each of the actions to be:**

(Mark with a tick ✓ in the box; one check for each row.)

Actions:	Very difficult	Difficult	Easy	Very easy	I don't know
25.1 Assessing whether a claim about a treatment is based on a research study comparing treatments (a fair comparison)	☐	☐	☐	☐	☐
25.2 Assessing where I can find information about treatments that is based on research studies comparing treatments (fair comparisons)	☐	☐	☐	☐	☐
25.3 Assessing how sure I can be about the results of a research study comparing treatments (the trustworthiness of the results)	☐	☐	☐	☐	☐
25.4 Assessing if the results of a research study comparing treatments are likely to be relevant to me	☐	☐	☐	☐	☐

26. Think about an illness that you might get. **How likely are you to say "yes" if you are asked to participate in a research study comparing two treatments for your illness (a fair comparison)?**				
(Mark with a tick ✓ in one box)				
Very unlikely	**Unlikely**	**Likely**	**Very likely**	**I don't know**
☐	☐	☐	☐	☐

Instructions: For the following questions, choose what you think is the best answer and write the letter for that answer in the box provided.

27. It is common for people to say that something will help improve your health or that it will not help. Some may say that it will be bad for your health. What people say about treatments may be correct but sometimes it may be wrong. We call these treatment claims.

How often do you hear treatment claims?

Options:

A) One treatment claim or more on most days

B) One treatment claim or more during most weeks

C) One treatment claim or more during most months

D) I almost never hear treatment claims

E) I don't know

Answer: ☐

28.1 When was the last time you heard a treatment claim?

Options:

A) This week

B) Last week

C) Last month

D) More than a month ago

E) I have never heard of any treatment claims

Answer: ☐

28.2 A treatment is anything done to care for yourself, so you stay well or, if you are sick or injured, so you get better and not worse

What was the **treatment** in the claim you last heard about?

Options:

A) Using a medicine (for example, taking a tablet or a syrup)

B) Getting an operation (for example, removing a bad tooth)

C) Using something to feel better or to heal more quickly (for example, using a bandage, or glasses)

D) Eating food or drinking something to feel better (for example, herbs or fruit)

E) Avoiding doing something to feel better (for example, not drinking milk)

F) Something else

Answer: ☐

28.3 In the space below please write down the treatment claim that you last heard. (What did they say the treatment would change or not change about someone's health?)

____

____

____

____

____

28.4 Did you think about what that treatment claim that you heard was based on?

Options:

- **A)** No, I did not think about what the treatment claim was based on
- **B)** Yes, I thought about what the treatment claim was based on
- **C)** I don't remember thinking about it

Answer: ☐

28.5 If you heard about a treatment claim, what was it based on?

It was based on:

- **A)** Someone's personal experience using the treatment
- **B)** What an expert said about it
- **C)** A research study that compared the treatment with another treatment or no treatment
- **D)** Something else
- **E)** I could not tell what the treatment claim was based on

Answer: ☐

28.6 How sure are you that the treatment claim you heard is true or can be trusted?

Options:

- **A)** Not very sure because I don't know the reason behind the claim
- **B)** Not very sure because the reason behind the claim was not good
- **C)** Very sure because the reason behind the claim was good
- **D)** I don't know because I don't know how to decide whether it is true or not

Answer: ☐

29.1 When was the last time you yourself decided whether to use a treatment or not to use a treatment?

Options:

- **A)** This week
- **B)** Last week
- **C)** Last month
- **D)** More than a month ago
- **E)** I have never decided to use or not to use a treatment

Answer: ☐

29.2 When was the last time that an adult decided for you whether you should use a treatment or not use a treatment?

Options:

- **A)** This week
- **B)** Last week
- **C)** Last month
- **D)** More than a month ago
- **E)** An adult has never decided for me
- **F)** I can't remember

Answer: ☐

29.3 What was the treatment for which you or an adult made the decision?

Options:

A) Using a medicine (for example, taking a tablet or a syrup)
B) Getting an operation (for example, removing a bad tooth)
C) Using something to feel better or to heal more quickly (for example, using a bandage, or glasses)
D) Eating food or drinking something to feel better (for example, herbs or fruit)
E) Avoiding doing something to feel better (for example, not drinking milk)
F) Something else

Answer: ☐

29.4 It is common for people to say that something will help improve your health or that it will not help. Some may say that it will be bad for your health. What people say may be correct but sometimes it may be wrong. We call these treatment claims.

What was the claim about the treatment for which you or an adult made the decision? (What did they say the treatment would change or not change about your health?)

____

____

____

29.5 Did you think about the possible reasons behind what they said about what that treatment would do or not do?

Options:

- **A)** No. I did not think about the possible reasons behind what they said
- **B)** Yes. I thought about the possible reasons behind what they said
- **C)** I don't remember thinking about the possible reasons behind what they said

Answer: ☐

29.6 What were the reasons behind what they said the treatment would change about your health?

Options:

- **A)** Someone's personal experience using the treatment
- **B)** What an expert said about it
- **C)** A research study that compared the treatment with another treatment or no treatment
- **D)** Something else
- **E)** I don't remember what the treatment claim was based on

Answer: ☐

29.7 What did you yourself decide to do about the treatment?

- **A)** I did not use the treatment
- **B)** I used the treatment
- **C)** I waited because I wanted to know more about the treatment
- **D)** My parents, or another person decided for me

Answer:

29.8 How sure are you about the advantages and disadvantages of the treatment you used?

Options:

- **A)** Not very sure because I don't know the reasons behind the claims about the good and bad things that treatment makes happen
- **B)** Not very sure because there was not a good reason behind the claims about the advantages of the treatment
- **C)** Not very sure because I only know about the advantages of the treatment. I also need to know about the disadvantages
- **D)** Very sure because there is a good reason behind the claims about the advantages and disadvantages of the treatment
- **E)** I did not use any treatment

Answer: ?

30. Who do you think should decide for you whether you should use a treatment or not use a treatment?

Options:

A) Me alone
B) My parents alone
C) Me and my parents
D) Me and someone in my family
E) A person who knows a lot about treatments
F) Me and a person who knows a lot about treatments
G) Me, my parents and a person who knows a lot about treatments

Answer:

Additional file 2. Additional tables

Table S1. Comparisons related to self-reported behaviours in the one-year follow-up

Question	Hypothesis and basis for the hypothesis
How often do you hear treatment claims?	Children in the intervention group will report hearing treatment claims more often because of being more aware of treatment claims and identifying them when they are made.
[For the last treatment claim that you heard,] did you think about what that treatment claim that you heard was based on?	A larger proportion of children in the intervention group will answer yes because of being more aware that many claims do not have a reliable basis.
How sure are you that the treatment claim you heard is true or can be trusted?	A smaller proportion of children in the intervention group will answer "very sure" or "I don't know", and a larger proportion of children in the intervention group will answer this question consistently with their answer to the preceding question about the basis of the claim (Table 3) because of being better able to assess the trustworthiness of claims and many claims not having a reliable basis.
How sure are you about the advantages and disadvantages of the [most recent] treatment you used?	A higher proportion of the children in the intervention group will answer "not very sure because I only know about the advantages" and a smaller proportion will answer "very sure", because information about the disadvantages of treatments is often lacking. However, this difference, if there is one, will likely be small, because children in the intervention group are more likely to consider and seek information about the disadvantages of treatments.
Who do you think should decide for you whether you should use a treatment or not use a treatment?	A higher proportion of the children in the intervention group will answer that they want to be included (A, C, D, F, or G) because of having learned about how to make informed health choices; and that someone who knows a lot about treatments should be included (E, F, or G), because of being more aware of the importance of assessing the reliability of evidence of effects and the skills that are needed to do this. However, this difference, if there is one, will likely be small, because children in the intervention group are more likely to recognise that expert opinion alone is not a reliable basis for a claim about treatment effects. What happens if the claim that comes in is about negative effects of the treatment? A larger proportion of children in the intervention group will answer, "Not very sure because there was not a good reason behind the claims about the advantages of the treatment", because they are more likely to identify a claim whose basis was bad.
Given your thoughts about the basis of the claim, what did you yourself decide to do about the treatment?	A smaller proportion of children in the intervention group versus the control group would choose to use a treatment (in question 29.7) having recognised that the basis of the claim was untrustworthy (in question 29.6)

Table S2. Ranges of marks and points awarded for each subject

Exam score (out of 100)	**Points awarded**	**Marks**
80-100	1	Distinction 1
70-79	2	Distinction 2
65-69	3	Credit 3
60-64	4	Credit 4
55-59	5	Credit 5
50-54	6	Credit 6
45-49	7	Pass 7
35-44	8	Pass 8
Below 35	9	Failure

Table S3. Exclusion criteria for self-reported behaviours

Response options for questions 28.2 and 29.3	**Response to questions 28.3 and 29.4**
28.2 What was the treatment in the claim you last heard about	28.3 Please write the claim that you last heard
29.3 What was the treatment for which you or an adult made the decision?	What was the claim about the treatment for which you or an adult made the decision?
Using a medicine (e.g. taking a tablet or syrup)	Exclude if the claim is not about a medicine
Getting an operation (e.g. removing a bad tooth)	Exclude if the claim is not about an operation
Using something to feel better or to heal more quickly (e.g. using a bandage or glasses)	Exclude if the claim is not about equipment
Something else (Eating food or drinking something to feel better (e.g. herbs or fruit))	Exclude if the claim is not about eating/drinking something e.g. herbs or fruit
Avoiding doing something to feel better (e.g. not drinking milk)	Exclude if the claim is not about avoiding something
Something else	Exclude if the claim is not about a treatment ("anything done to care for yourself, so you stay well or, if you are sick or injured, so you get better and not worse")

Table S4. Number of missing values for each question

	Number of unanswered questions			
	Control (n=2844)		Intervention (n=3943)	
	N	%	N	%
Question 4	9	0.32%	20	0.51%
Question 5	8	0.28%	26	0.66%
Question 6	8	0.28%	17	0.43%
Question 7	8	0.28%	16	0.41%
Question 8	8	0.28%	25	0.63%
Question 9	10	0.35%	16	0.41%
Question 10	3	0.11%	13	0.33%
Question 11	7	0.25%	7	0.18%
Question 12	12	0.42%	21	0.53%
Question 13	8	0.28%	19	0.48%
Question 14	24	0.84%	38	0.96%
Question 15	22	0.77%	39	0.99%
Question 16	19	0.67%	30	0.76%
Question 17	31	1.09%	35	0.89%
Question 18	21	0.74%	16	0.41%
Question 19	19	0.67%	16	0.41%
Question 20	32	1.13%	19	0.48%
Question 21 part 1	79	2.78%	39	0.99%
Question 21 part 2	96	3.38%	54	1.37%
Question 21 part 3	95	3.34%	53	1.34%
Question 21 part 4	87	3.06%	43	1.09%
Question 22 part 1	79	2.78%	50	1.27%
Question 22 part 2	59	2.07%	54	1.37%
Question 22 part 3	93	3.27%	70	1.78%
Question 23 part 1	69	2.43%	53	1.34%
Question 23 part 2	52	1.83%	44	1.12%
Question 24 part 1	69	2.43%	67	1.70%
Question 24 part 2	87	3.06%	70	1.78%
Question 25 part 1	71	2.50%	55	1.39%
Question 25 part 2	73	2.57%	71	1.80%
Question 25 part 3	82	2.88%	84	2.13%
Question 25 part 4	72	2.53%	86	2.18%

Table S5. Attendance and national examinations

	Control schools	Intervention schools	Adjusted difference	P-value
Attendance rates				
	Control schools N=33 schools Median (25th to 75th percentile)	**Intervention schools** N=31 schools Median (25th to 75th percentile)	**Adjusted difference**	**P-value**
Intervention term	90.3% (78.7% to 98.0%)	89.1% (80.4% to 96.4%)	3% less (95% CI -14 to 6)	0.437
Following term	91.7% (81.1% to 97.8%)	89.5% (78.6% to 96.2%)	2% more (95% CI -10 to 13)	0.726
Average scores on national examinations				
	Control schools Mean (SD)	**Intervention schools** Mean (SD)	**Adjusted mean difference**	**P-value**
End of intervention term				
English	54.2% (22.5)	52.3% (22.5)	-1.7% (95% CI -6.6 to 3.2)	0.500
Math	51.5% (23.4)	49.0% (22.5)	-1.8% (95% CI -6.6 to 3.0)	0.457
Science	49.8% (24.4)	49.7% (23.3)	-0.5% (95% CI -5.4 to 4.5)	0.852
Social science	52.6% (24.0)	51.9% (23.7)	-1.0% (95% CI -6.2 to 4.2)	0.699
Total	52.3% (21.4)	51.1% (21.0)	-1.2% (-5.5 to 3.2)	0.597
Following term				
English	56.3% (22.1)	56.1% (22.5)	2.4% (95% CI -2.3 to 7.2)	0.312
Math	53.8% (23.2)	50.2% (22.4)	0.8% (95% CI -4.1 to 5.8)	0.752
Science	52.4% (23.9)	49.3% (23.3)	0.8% (95% CI -4.1 to 5.4)	0.813
Social science	56.0% (23.8)	52.0% (22.7)	-0.1% (95% CI -4.8 to 4.7)	0.964
Total	54.8% (21.5)	52.2 % (20.6)	1.0% (-3.4, 5.4)	0.671
Proportion with a passing score (≥ 35%) on the national examinations				
	Control schools N (%)	**Intervention schools** N (%)	**Adjusted difference**	
End of intervention term	Total: 49 schools, 3795 children	Total: 44 schools, 4201 children		
English	2917/3766 (77.5%)	3009/3984 (71.8%)	0.0% (95% CI -10.0 to 13.8)	0.998
Math	2709/3772 (71.8%)	2809/3985 (70.5%)	1.6% (95% CI -12.0 to 11.9)	0.799
Science	2632/3764 (69.9%)	2829/3990 (70.9%)	-0.1% (95% CI -11.4 to 14.6)	0.988
Social science	2794/3773 (74.1%)	2957/3980 (74.3%)	-1.7% (95% CI -11.9 to 12.9)	0.801
Total	2698/3730 (72.3%)	2830/3934 (71.9%)	-0.7% (95% CI -11.5 to 13.8)	0.920
Following term	Total: 51 schools, 3956 children	Total: 48 schools, 4474 children		
English	3205/3934 (81.5%)	3655/4460 (82.0%)	3.8% (95% CI -5.2 to 16.6)	0.461
Math	3038/3940 (76.9%)	3174/4441 (71.5%)	-0.1% (95% CI -10.3 to 12.8)	0.984
Science	2923/3942 (74.2%)	3137/4436 (70.7%)	-0.1% (95% CI -11.4 to 14.6)	0.878
Social science	3125/3940 (79.3%)	3366/4452 (75.6%)	1.1 (95% CI -8.1 to 13.2)	0.839
Total	3022/3914 (77.2%)	3268/4404 (74.2%)	1.5% (95% CI -8.6 to 14.8)	0.797
Proportion with a distinction score (≥ 70%) on the national examinations				
	Control schools N (%)	**Intervention schools** N (%)	**Adjusted difference**	
End of intervention term	Total: 49 schools, 3795 children	Total: 44 schools, 4201 children		
English	1133/3766 (30.1%)	1077/3984 (27.0%)	-7.0% (95% CI -21.4 to 4.9)	0.278
Math	995/3772 (26.4%)	850/3985 (21.3%)	-4.2% (95% CI -17.3 to 5.6)	0.716
Science	966/3764 (25.7%)	977/3990 (24.5%)	-2.1% (95% CI -14.9 to 7.7)	0.716
Social science	1117/3773 (29.6%)	1117/3980 (28.1%)	-1.7% (95% CI -15.5 to 9.2)	0.791
Total	904/3730 (24.2%)	882/3934 (22.4%)	-2.1% (95% CI-15.0 to 7.3)	0.693
Following term	Total: 51 schools, 3956 children	Total: 48 schools, 4474 children		
English	1263/3934 (32.1%)	1440/4460 (32.3%)	4.8% (95% CI -7.7 to 14.6)	0.425
Math	1101/3940 (27.9%)	1023/4441 (23.0%)	-3.4% (95% CI -16.8 to 6.6)	0.551
Science	1099/3942 (27.9%)	1024/4436 (23.1%)	-0.8% (95% CI -12.3 to 7.9)	0.875
Social science	1342/3940 (34.1%)	1207/4452 (27.1%)	-0.2% (95% CI -12.4 to 9.3)	0.967
Total	1063 (27.2%)	1012 (23.0%)	1.3% (95%CI -11.1 to 10.0)	0.819

SD = standard deviation

Table S6. Sensitivity analyses - one-year follow-up

	Adjusted difference*	Odds ratio
Mean score		
Primary analysis	**Mean difference: 16.7%** (95% CI 13.9% to 19.5%) P <0.00001	
Weighted analysis	**Mean difference: 16.7%** (95% CI 13.9% to 19.5%)	
Lee bounds	**6.4% to 26.6%** (95% CI 6.6% to 26.5%)	
Passing score (≥ 13 out of 24 correct answers)		
Primary analysis	**39.5%** (95% CI 29.9%-47.5%)	**5.88** (95% CI 4.00 to 8.33) P<0.0001
Weighted analysis	**40.9%** (95% CI 31.0% to 49.4%)	**6.25** (95% CI 4.17 to 9.09) P<0.0001

* The adjusted difference is based on mixed models with a random effects term for the clusters and the stratification variables modelled as fixed effects, using logistic regression for dichotomous outcomes and linear regression for continuous outcomes. The odds ratios from the logistic regressions for passing scores have been converted to differences based on the intervention school proportions and the odds ratios calculated using the intervention schools as the reference (the inverse of the odds ratios shown here).

Table S7. Attrition, differences in test scores across strata of schools

Proportion of children who completed the test

Control schools

	Government	**Private**	**Total**	**N (%) of schools**
Rural	35.1% (124/353)	43.7% (62/142)	37.6% (186/495)	8 (13.3%)
Semi-urban	40.8% (430/1055)	53.7% (116/216)	43% (546/1271)	15 (25%)
Urban	43.1% (957/2219)	50.9% (1155/2271)	47% (2112/4490)	37 (61.7%)
Total	**41.7% (1511/3627)**	**50.7% (1333/2629)**	**45.5% (2844/6256)**	**60 (100%)**

Intervention schools

	Government	**Private**	**Total**	**N (%) of schools**
Rural	42.1% (117/278)	60.6% (175/289)	51.5% (292/567)	6 (10%)
Semi-urban	59.6% (766/1286)	75.7% (467/617)	64.8% (1233/1903)	14 (23.3%)
Urban	60.6% (1406/2319)	63.5% (1012/1594)	61.8% (2418/3913)	40 (66.7%)
Total	**58.9% (2289/3883)**	**66.2% (1654/2500)**	**61.8% (3943/6383)**	**60 (100%)**

Test scores

	Treatment effect	School ownership effect	School location effect
Mean score			
Without weighting	16.7% (95% CI 13.9 to 19.5) P < 0.00001	7.2% (95% CI 4.3 to 10.0) P < 0.00001	0.2% (95% CI -1.9 to 2.3) P = 0.872
Weighted	16.7% (95% CI 13.9 to 19.5) P < 0.00001	7.2% (95% CI: 4.3 to 10.1) P < 0.00001	0.3% (95% CI -1.8to 2.3) P = 0.807
Passing score (≥ 13 out of 24 correct answers)			
Without weighting	OR 0.17 (95% CI 0.12 to 0.25) P < 0.00001	OR 2.14 (95% CI 1.49 to 3.09) P = 0.00004	OR 0.99 (95% CI 0.76 to 1.29) P = 0.92
Weighted	OR 0.16 (95% CI 0.11 to 0.24) P < 0.00001	OR 2.28 (95% CI 1.54 to 3.38) P = 0.00004	OR 1.00 (95% CI 0.75 to 1.32) P = 0.99
Mastery score (≥ 20 out of 24 correct answers)			
Without weighting	OR 0.10 (95% CI 0.06 to 0.15) P < 0.00001	OR 2.34 (95% CI 1.59 to 3.46) P = 0.00002	OR 0.99 (95% CI 0.74 to 1.33) P = 0.951
Weighted	OR 0.09 (95% CI 0.06 to 0.13) P < 0.00001	OR 2.59 (95% CI 1.72 to 3.90) P = 0.0005	OR 1.06 (95% CI 0.78 to 1.44) P = 0.712

OR: odds ratio

Table S8. Intended behaviours - one-year follow-up

Think about an illness that you might get. Imagine someone claiming (saying) that a particular treatment might help you get better.

	How likely are you to find out what the claim was based on (for example by asking the person making the claim)?		How likely are you to find out if the claim was based on a research study comparing the treatment to no treatment (a fair comparison)?		How likely are you to say "yes" if you are asked to participate in a research study comparing two treatments for your illness (a fair comparison)?	
	Control schools N=2844	**Intervention schools** N=3943	**Control schools** N=2844	**Intervention schools** N=3943	**Control schools** N= 2844	**Intervention schools** N= 3943
Missing	69 (2.4%)	67 (1.7%)	87 (3.1%)	70 (1.8%)	36 (1.3%)	44 (1.1%)
Very unlikely	217 (7.6%)	376 (9.5%)	301 (10.6%)	467 (11.8%)	245 (8.6%)	277 (7.0%)
Unlikely	289 (10.2%)	376 (9.5%)	424 (14.9%)	569 (14.4%)	329 (11.6%)	429 (10.9%)
Likely	975 (34.3%)	1510 (38.3%)	747 (26.3%)	997 (25.3%)	1045 (36.7%)	1577 (40.0%)
Very likely	678 (23.8%)	1082 (27.4%)	705 (24.8%)	1164 (29.5%)	719 (25.3%)	1155 (29.3%)
I don't know	616 (21.7%)	532 (13.5%)	580 (20.4%)	676 (17.1%)	470 (16.5%)	461 (11.7%)
Likely or very likely*	**1653 (58.1%)**	**2592 (65.7%)**	**1452 (51.1%)**	**2161 (54.8%)**	**1764 (62.0%)**	**2732 (69.3%)**
Odds ratio (95% CI)†	1.41 (1.18 - 1.69) P=0.00020		1.11 (0.93 - 1.33) P=0.269		1.41 (1.10 - 1.79) P=0.00629	
Adjusted Difference†	**8.1%** (3.7%-12.6%)		**2.6%** (-1.9% - 7.2%)		**7.7%** (2.0% - 13.5%)	
End of intervention term‡						
Likely or very likely	**2440 (55.1%)**	**3731 (64.9%)**	**1967 (44.4%)**	**3114 (54.1%)**	**2163 (48.8%)**	**3201 (55.6%)**
Odds ratio	1.56 (95% CI 1.29 to 1.88)		1.54 (95% CI 1.29 to 1.84)		1.37 (95% CI 1.16 to 1.62)	
Adjusted Difference	**10.6%** (95% CI 6.2% to 14.7%)		**10.8%** (95% CI 6.3% to 15.1%)		**7.8%** (95% CI 3.7% to 11.9%)	

* Missing values and don't know are pooled with unlikely and very unlikely.

† The difference is an adjusted difference, based on mixed models with a random effects term for the clusters and the stratification variables modelled as fixed effects, using logistic regression. The odds ratios from the logistic regressions have been converted to differences using the intervention schools as the reference and the inverse of the odds ratios shown here.

‡ Results based on responses at the end of the term when the intervention was delivered.

Table S9. Self-efficacy

How difficult or easy would you find each of these actions to be?

	Assessing whether a claim about a treatment is based on a research study comparing treatments (a fair comparison)		Assessing where I can find information about treatments that is based on research studies comparing treatments (fair comparisons)		Assessing how sure I can be about the results of a research study comparing treatments (the trustworthiness of the results)		Assessing if the results of a research study comparing treatments are likely to be relevant to me	
	Control schools N=2844	**Intervention schools** N=3943	**Control schools** N=2844	**Intervention schools** N=3943	**Control schools** N=2844	**Intervention schools** N=3943	**Control schools** N=2844	**Intervention schools** N=3943
Missing	71 (2.5%)	55 (1.4%)	73 (2.6%)	71 (1.8%)	82 (2.9%)	84 (2.1%)	72 (2.5%)	86 (2.2%)
Very difficult	357 (12.6%)	455 (11.5%)	338 (11.9%)	431 (10.9%)	488 (17.2%)	581 (14.7%)	436 (15.3%)	568 (14.4%)
Difficult	779 (27.4%)	865 (21.9%)	634 (22.3%)	876 (22.2%)	653 (23.0%)	1007 (25.5%)	513 (18.0%)	727 (18.4%)
Easy	837 (29.4%)	1517 (38.5%)	899 (31.6%)	1348 (34.2%)	640 (22.5%)	897 (22.7%)	694 (24.4%)	1027 (26.0%)
Very easy	334 (11.7%)	623 (15.8%)	525 (18.5%)	856 (21.7%)	454 (16.0%)	712 (18.1%)	562 (19.8%)	779 (19.8%)
I don't know	466 (16.4%)	428 (10.9%)	375 (13.2%)	361 (9.2%)	527 (18.5%)	662 (16.8%)	567 (19.9%)	756 (19.2%)
Easy or very easy*	**1171 (41.2%)**	**2140 (54.3%)**	**1424 (50.1%)**	**2204 (55.9%)**	**1094 (38.5%)**	**1609 (40.8%)**	**1256 (44.2%)**	**1806 (45.8%)**
Odds ratio (95% CI)†	1.82 (1.43 - 2.33) P<0.00001		1.33 (1.11 - 1.59) P=0.00171		1.10 (0.94 - 1.30) P=0.233		1.10 (0.93 - 1.28) P=0.279	
Adjusted difference†	**14.8%** (8.9% - 20.5%)		**7.2%** (2.6% – 11.5%)		**2.3%** (-1.4% - 6.1%)		**2.3%** (-1.9% - 6.1%)	
End of intervention term‡								
Easy or very easy	**1886 (42.6%)**	**3244 (56.4%)**	**3069 (53.3%)**	**2238 (50.5%)**	**1777 (40.1%)**	**2112 (36.7%)**	**2002 (45.2%)**	**2727 (47.4%)**
Odds ratio	1.83 (95% CI 1.55 to 2.16)		1.13 (95% CI 0.96 to 1.33)		0.84 (95% CI 0.73 to 0.96)		1.08 (95% CI 0.93 to 1.25)	
Adjusted difference	**15.0%** (95% CI 10.9% to 19.0%)		**3.0%** (95% CI -1.0% to 7.0%)		**-4.1%** (95% CI -1.0% to -7.3%)		**1.9%** (95% CI -1.8% to 5.6%)	

* Missing values and don't know are pooled with difficult and very difficult.

† The difference is an adjusted difference, based on mixed models with a random effects term for the clusters and the stratification variables modelled as fixed effects, using logistic regression. The odds ratios from the logistic regressions have been converted to differences using the intervention schools as the reference and the inverse of the odds ratios shown here.

‡ Results based on responses at the end of the term when the intervention was delivered.

Table S10. Self-reported behaviour - awareness of treatment claims

How often do you hear treatment claims?

	Control schools N=2844	Intervention schools N=3943
One or more most days	572 (20.1%)	1000 (25.4%)
One or more most weeks	374 (13.2%)	599 (15.2%)
One or more most months	497 (17.5%)	715 (18.1%)
Almost never	653 (23.0%)	788 (20.0%)
I don't know	717 (25.2%)	810 (20.5%)
Missing	31 (1.1%)	31 (0.8%)
One or more most days or most weeks	**946 (33.8%)**	**1599 (40.6%)**
Odds ratio*	1.35 (95% CI 1.02 - 1.79) P = 0.0356	
Adjusted difference†	**7.0%** (95% CI 0.5% to 12.9%)	

*The odds ratio for the dichotomised data is shown in the table. The odds ratio from the mixed ordinal logistic regression was 1.30 (95% CI 1.01 to 1.67, P = 0.0431).

† The difference is an adjusted difference, based on a mixed model with a random effects term for the clusters and the stratification variables modelled as fixed effects, using logistic regression. The odds ratio from the logistic regression has been converted to a difference using the intervention schools as the reference and the inverse of the odds ratios shown here.

Table S11. Self-reported behaviour - assessment of trustworthiness of treatment claims

How sure are you that the treatment claim you heard is true or can be trusted?

	Control schools N=2844	**Intervention schools** N=3943
Missing	49 (1.7%)	60 (1.5%)
Not very sure because I don't know the reason behind the claim	665 (23.4%)	1039 (26.4%)
Not very sure because the reason behind the claim was not good	543 (19.1%)	1087 (27.6%)
Very sure because the reason behind the claim was good	704 (24.8%)	790 (20.0%)
I don't know because I don't know how to decide whether it is true or not	883 (31.0%)	967 (24.5%)
Very sure or I don't know	**1587 (55.8%)**	**1757 (44.6%)**
Odds ratio (very sure or I don't know vs other)	0.55 (95% CI 0.45 - 0.67) P<0.0001	
Adjusted difference*	**-15.0%** (95% CI -9.9% to -19.7%)	
Odds ratio (consistent with what they identified as the basis for the claim)†	1.45 (95% CI 1.18 - 1.75) P=0.000549	
Adjusted difference*	**7.6%** (95% CI 3.5% - 11.1%)	

* The differences are adjusted differences, based on mixed models with a random effects term for the clusters and the stratification variables modelled as fixed effects, using logistic regression. The odds ratio from the logistic regression has been converted to a difference using the intervention schools as the reference and the inverse of the odds ratios shown here.

† See Table S11.

Table S12. Consistent (correct) answers regarding certainty about treatment claims*

If you heard about a treatment claim, what was it based on?	How sure are you that the treatment claim you heard is true or can be trusted?
Someone's personal experience using the treatment	Not very sure because the reason behind the claim was not good
What an expert said about it	Not very sure because the reason behind the claim was not good
A research study that compared the treatment with another treatment or no treatment	Not very sure because the reason behind the claim was not good OR Very sure because the reason behind the claim was good
Something else	Not very sure because the reason behind the claim was not good
I could not tell what the treatment claim was based on	Not very sure because I don't know the reason behind the claim

* Questions 28.5 and 28.6 in Appendix 1

Table S13. Self-reported behaviour - assessment of the basis of treatment claims

For the last treatment claim that you heard, did you think about what that treatment claim that you heard was based on?

	Control schools N=2844	**Intervention schools** N=3943
Missing	50 (1.8%)	57 (1.4%)
No	512 (18.0%)	845 (21.4%)
Yes	1387 (48.8%)	2116 (53.7%)
I don't remember	895 (31.5%)	925 (23.5%)
Odds ratio (yes versus other)	1.18 (95% CI 0.95 - 1.47) P=0.130	
Adjusted difference*	**4.1%** (95% CI -1.2% - 9.6%)	

* The difference is an adjusted difference, based on a mixed model with a random effects term for the clusters and the stratification variables modelled as fixed effects, using logistic regression. The odds ratio from the logistic regression has been converted to a difference using the intervention schools as the reference and the inverse of the odds ratios shown here.

Table S14. Self-reported behaviour - assessment of advantages and disadvantages of treatments

How sure are you about the advantages and disadvantages of the [most recent] treatment you used?

	Control schools N=2844	**Intervention schools** N=3943
A) Not very sure because I don't know the reasons behind the claims about the good and bad things that treatment makes happen	531 (18.7%)	851 (21.6%)
B) Not very sure because there was not a good reason behind the claims about the advantages of the treatment	355 (12.5%)	549 (13.9%)
C) Not very sure because I only know about the advantages of the treatment. I also need to know about the disadvantages	765 (26.9%)	992 (25.2%)
D) Very sure because there is a good reason behind the claims about the advantages and disadvantages of the treatment	652 (22.9%)	929 (23.6%)
E) I did not use any treatment	498 (17.5%)	590 (15.0%)
Missing	43 (1.5%)	32 (0.8%)
Odds ratio (C versus any other response)	1.05 (95% CI 0.86 - 1.30) P=0.62	
Adjusted difference answer C vs else	**-0.9%** (95% CI -5.3% - 2.7%)	
Odds ratio (D versus any other response)	1.03 (95% CI 0.85 - 1.23) P=0.79	
Adjusted difference answer D vs else	**-0.5%** (95% CI -3.9% - 2.8%)	

Table S15. Subgroup analysis - reading skills

	Control schools	Intervention schools	Adjusted difference†	Odds ratio	ICC
Mean score, %					
Lacking basic reading skills (N=1775)	N children = 893 Mean score: 47.2% (SD 16.4%)	N children = 882 Mean score: 57.1% (SD 18.1%)	**Mean difference: 11.2%** (95% CI 8.2% to 14.2%)		0.146
Basic reading skills (N=2672)	N children = 1093 Mean score: 55.2% (SD 16.9%)	N children = 1579 Mean score: 67.9% (SD 16.8%)	**Mean difference: 14.8%** (95% CI 12.3% to 17.3%)		0.162
Advanced reading skills (N=2340)	N children = 858 Mean score: 56.3% (SD 15.6%)	N children = 1482 Mean score: 76.5% (SD 15.5%)	**Mean difference: 19.4%** (95% CI 16.9% to 21.9%)		0.117
Passing score ($\geq$ 13 out of 24 correct answers)					
Lacking basic reading skills (N=1775)	N children = 893 36.6% of children N=327	N children = 882 59.3% of children N=523	**28.9% more children** (95% CI 20.8% to 36.7%)	0.30 (95% CI 0.20 to 0.43)	0.144
Basic reading skills (N=2672)	N children = 1093 57.0% of children N=623	N children = 1579 81.2% of children N= 1282	**33.6% more children** (95% CI 24.0% to 41.9%)	0.21 (95% CI 0.15 to 0.31)	0.150
Advanced reading skills (N=2340)	N children = 858 60.0% of children N=514	N children = 1482 91.4% of children N=1355	**33.4% more children** (95% CI 25.7% to 42.5%)	0.13 (95% CI 0.09 to 0.18)	0.098
Mastery score ($\geq$ 20 out of 24 correct answers)					
Lacking basic reading skills (N=1775)	N children = 893 3.0 % of children N=27	N children = 882 10,1 % of children N=89	**7.7% more children** (95% CI 5.6% to 8.8%)	0.22 (95% CI 0.12 to 0.42)	0.220
Basic reading skills (n=2672)	N children = 1093 6.5% of children N=71	N children = 1579 24.1% of children N=380	**19.6% more children** (95% CI 17.0% to 21.3%)	0.15 (95% CI 0.09 to 0.24)	0.192
Advanced reading skills (n=2340)	N children = 858 4.8% of children N=41	N children = 1482 45.1% of children N=669	**40.4% more children** (95% CI 38.2% to 41.9%)	0.06 (95% CI 0.04 to 0.09)	0.139

* Because reading skills were measured after the intervention, we have not reported a test of interaction here (see Appendix 3).

† The adjusted difference is based on mixed models with a random effects term for the clusters and the stratification variables modelled as fixed effects, using logistic regression for dichotomous outcomes and linear regression for continuous outcomes. The odds ratios from the logistic regressions for passing scores and mastery scores have been converted to differences using the intervention school proportions and the inverse of the odds ratios shown here.

Table S16. Differences in reading skills

Reading skills	Immediately after the intervention*			One-year follow-up*			Change from first to second test*		
	Control schools N children 4412 n (%)	Intervention schools N children 5711 n (%)	Diff	Control schools N children 2844 n (%)	Intervention schools N children 3943 n (%)	Diff	Control schools	Intervention schools	Diff
Lacking basic reading skills	2139 (48.5%)	2224 (38.9%)	-9.5%	893 (31.4%)	882 (22.4%)	-9.0%	-17.1%	-16.6%	0.5%
Basic reading skills	1507 (34.2%)	2155 37.7%	3.6%	1093 (38.4%)	1579 (40.0%)	1.6%	4.3%	2.3%	-2.0%
Advanced reading skills	766 (17.4%)	1332 23.3%	6.0%	858 (30.2%)	1482 (37.6%)	7.4%	12.8%	14.3%	1.5%

* Reading skills as measured by first four questions in the test administered at the end of the term when the intervention was delivered and the same test one year later. The differences (Diff) are shown between the intervention and control schools for each time the test was administered and the change from the first to the second time.

Table S17. Subgroup analysis - parent who listened to the podcast

	Control schools	Intervention schools	Adjusted effect of the interaction*
	N children = 69	N children = 98	
Parent in control group (N=167)	Mean score: 55.1% (SD 16.4%)	Mean score: 64.5% (SD 20.2%)	**Mean difference: 3.8%** (95% CI -3.9% to 11.4%) P=0.3443
	N children = 64	N children = 104	
Parent in podcast group (N=168)	Mean score: 53.6% (SD 15.9%)	Mean score: 66.3% (SD 18.6%)	

*Adjusted for location, ownership and random effect of clustering, ICC=0.185

Table S18. Exploratory analyses - p-values for differences between first (end of intervention term) and second (one-year follow-up) effects

	Control schools	Intervention schools	Adjusted difference*	Odds ratio*
Primary outcome				
Mean score (%)				
1. End of intervention term	Mean score: 43.1% (SD 15.2%)	Mean score: 62.4% (SD 18.8%)	20.0% (95% CI 17.3% to 22.7%)	
2. One-year follow-up	Mean score: 53.0% (SD 16.8%)	Mean score: 68.7% (SD 18.2%)	16.7% (95% CI 13.9% to 19.5%)	
Difference between first and second tests			**-6.9% (95% CI 6.5 to 7.3) P<0.00001**	
Passing score†				
1. End of intervention term	26.8 % of children (N=1186/4430)	69.0 % of children (N=3967/5753)	49.8% more children (95% CI 43.8% to 54.6%)	9.34 (95% CI 6.62 to 13.18)
2. One-year follow-up	51.5 % of children (N=1464/2844)	80.1 % of children (N=3160/3943)	39.5% more children (95% CI 29.9% to 47.5%)	5.88 (95% CI 4.00 to 8.33)
Difference between first and second tests				**0.36 (95% CI: 0.33 to 0.39) P<0.00001**
Secondary outcomes				
Mastery score‡				
1. End of intervention term	0.9% of children (N=38/4430)	18.6% of children (N=1070/5753)	18.0% more children (95% CI 17.5% to 18.2%)	35.33 (95% CI 20.58 to 60.67)
2. One-year follow-up	4.9% of children (N=139/2844)	28.9% of children (N=1138/3943)	25.0% more children (95% CI 23.2% to 26.5%)	10.00 (95% CI 6.67 to 16.67)
Difference between first and second tests				**0.42 (95% CI: 0.37 to 0.47) P<0.00001**

* The differences between effects (adjusted mean differences and odds ratios) from is based on mixed models with a random effects term for the clusters (schools), individuals (who are used twice in these analyses), and the stratification variables modelled as fixed effects, using logistic regression for dichotomous outcomes and linear regression for continuous outcomes.

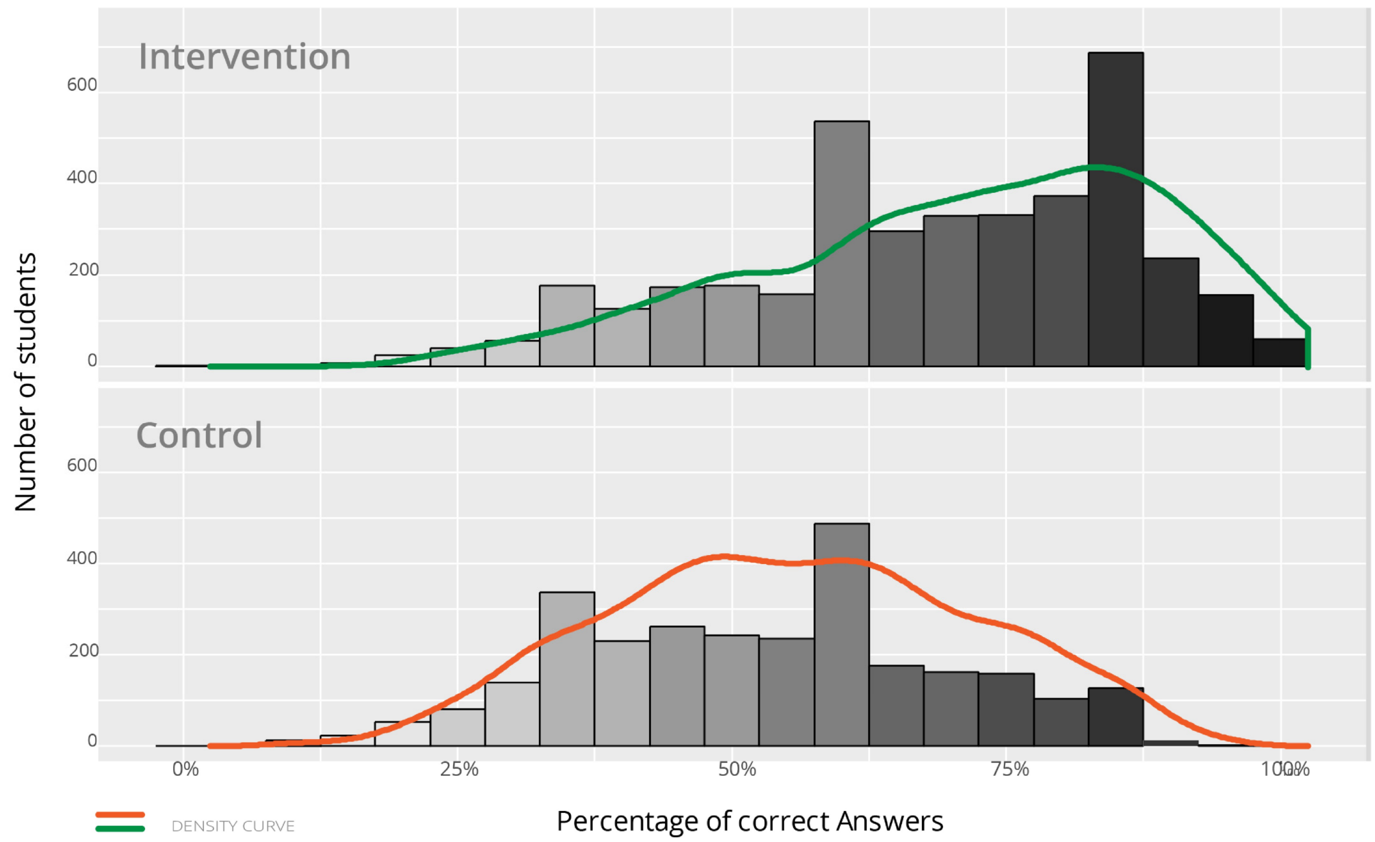
† 13 or more out of 24 correct answers

‡ 20 or more out of 24 correct answers

Table S19. Exploratory analyses excluding children who did not take the test both times

	Control schools	Intervention schools	Adjusted difference*	Odds ratio*	ICC
Primary outcome					
One-year follow-up Mean score, %	Mean score: 53.0% (SD 16.8%)	Mean score: 68.7% (SD 18.2%)	**Mean difference: 16.7%** (95% CI 13.9% to 19.5%) P <0.00001		0.18
End of intervention term excluding children who did not take the one-year follow-up test Mean score, %	Mean score: 43.8% (SD 15.5%) (N=2733)	Mean score: 64.6% (SD 18.5%) (N=3875)	Mean difference: 21.6% (95% CI 18.9 – 24.4) P<0.00001		0.17
End of intervention term Mean score, %	Mean score: 43.1% (SD 15.2%)	Mean score: 62.4% (SD 18.8%)	Mean difference: 20.0% (95% CI 17.3% to 22.7%)		0.18
One-year follow-up Passing score (≥ 13 out of 24 correct answers)	51.5 % of children (N=1464/2844)	80.1 % of children (N=3160/3943)	**39.5% more children** (95% CI 29.9% to 47.5%)	5.88 (95% CI 4.00 to 8.33) P <0.00001	0.20
End of intervention term excluding children who did not take the one-year follow-up test Passing score (≥ 13 out of 24 correct answers)	28.4% (776/2733)	74.0% 2867/(3875)	**54.1% (95% CI: 47.8 – 59.1)** **P<0.00001**	0.09 (0.06 – 0.12)	0.19
End of intervention term Passing score (≥ 13 out of 24 correct answers)	26.8 % of children (N=1186/4430)	69.0 % of children (N=3967/5753)	49.8% more children (95% CI 43.8% to 54.6%)	9.34 (95% CI 6.62 to 13.18)	0.19
Secondary outcomes					
One-year follow-up Mastery score (≥ 20 out of 24 correct answers)	4.9% of children (N=139/2844)	28.9% of children (N=1138/3943)	**Mean difference: 25.0%** **(23.2%-26.5%)**	10.00 (95% CI 6.67 to 16.67) P <0.00001	0.19
End of intervention term excluding children who did not take the one-year follow-up test Mastery score (≥ 20 out of 24 correct answers)	0.8% (N 23/2733)	21.9% (N 847/3875)	**21.2% (95% CI 20.7 – 21.5)**	0.02 (0.01 – 0.04) P<0.00001	0.16
End of intervention term Mastery score (≥ 20 out of 24 correct answers)	0.9% of children (N=38/4430)	18.6% of children (N=1070/5753)	**18.0% more children** (95% CI 17.5% to 18.2%)	35.33 (95% CI 20.58 to 60.67)	0.21

Follow-up one year after the intervention was delivered



End of term when the intervention was delivered