BMJ Open | Impact of an SMS advice programme on maternal and newborn health in rural China: study protocol for a quasi-randomised controlled trial

Yanfang Su,[1] Changzheng Yuan,[2] Zhongliang Zhou,[3] Jesse Heitner,[4] Benjamin Campbell[5]

CrossMark

For numbered affiliations see end of article.

Correspondence to
Professor Zhongliang Zhou;
zzliang1981@mail.xjtu.edu.cn
and Dr Changzheng Yuan;
Chy478@mail.harvard.edu

ABSTRACT

Introduction: Expectant mothers in low-income and middle-income countries often lack access to vital information about pregnancy, preparation for birth and best practices when caring for their newborn. Innovative solutions are needed to bridge this knowledge gap and dramatically improve maternal and neonatal health in these settings. This study aims to evaluate the impact of an innovative text messaging intervention on maternal and neonatal health outcomes.

Methods and analysis: This study offers expectant mothers in rural China a package of free short messages via cell phone regarding pregnancy and childbirth. These messages are tailored to each mother's gestational week. It is hypothesised that delivering these short advice messages to pregnant women can improve maternal and newborn health. The study uses factorial quasi-randomisation to compare psychological, behavioural and health outcomes between 4 groups: 2 groups receiving different sets of short message interventions (ie, good household prenatal practices and healthcare seeking), a group receiving both interventions and a control group. Treatment assignment occurs at the individual level. The primary outcome is newborn health, measured by appropriateness of weight for gestational age. Secondary outcomes include severe neonatal and maternal morbidity as well as psychological and behavioural measures. This study has enrolled pregnant women who attend county maternal and child health centres for their prenatal visits.

Discussion: This pilot is the first large-scale effort to build a comprehensive evidence base on the impact of prenatal text messages via cell phone on maternal and newborn health outcomes in China. The study has broad implications for public health policy in China and the implementation of mobile health interventions in low-resource settings around the world.

Ethics: This study was approved by the Ethics Committee of the School of Medicine at Xi'an Jiaotong University on 18 January 2013.

Trial registration number: NCT02037087;
Pre-results.

Strengths and limitations of this study

- This is the first large-scale effort to build a comprehensive evidence base on the impact of prenatal text messages via cell phone on maternal and newborn health outcomes in China.
- The primary outcome is newborn health and the secondary outcomes include severe neonatal and maternal morbidity as well as psychological and behavioural measures.
- The study has broad implications for public health policy in China and the implementation of mobile health interventions in low-resource settings around the world.
- We tried to create the 'placebo' by sending messages to the control group and we admit the difference in the number of messages by group is one of our limitations.

INTRODUCTION

Expectant mothers in low-income and middle-income countries often lack access to vital information about pregnancy, preparation for birth and best practices when caring for their newborn. Furthermore, traditional cultural routines sustain suboptimal prenatal practices. Innovative solutions are needed to bridge this knowledge gap and dramatically improve maternal and child health in low-resource settings. One potential solution is the use of mobile health (mHealth), and specifically the use of text messages (short message service, SMS), to accurately inform underserved women. However, according to systematic reviews examining the use of SMS messages,[1] [2] the impact of SMS on maternal and child health has not been evaluated in a sufficiently powered randomised controlled trial. This study was motivated to fill this important gap in the literature and to ultimately translate

research into practice by providing policy recommendations. It is hypothesised that the message intervention will have a psychological impact, which will then promote a behavioural impact, which will finally lead to health outcome improvement. The study aims to determine if and how an innovative SMS intervention providing educational information to pregnant women in China might lead to improved maternal care as well as maternal and child health outcomes.

LITERATURE REVIEW

The use of SMS has emerged as a powerful tool in public health and has shown to be effective in changing behaviour. For example, studies in the USA, Japan and the UK showed that SMS reminders led to positive outcomes in smoking cessation and physical activity promotion among young populations.[3][4] In addition, it has been demonstrated that SMS can be used as an effective outreach strategy to inform a targeted population about government-subsidised services and second opinions from an independent third party.[5][6] Furthermore, in low-resource settings, SMS has been shown to improve attendance to clinic appointments[7] and uptake of vaccinations.[8] The literature about the impact of SMS on behavioural changes supports that process and satisfaction outcomes can be improved by SMS[9–12] and this has led many to posit that clinical and health outcomes can be improved via SMS as well.

One highly promising avenue for using SMS to improve clinical and health outcomes is in the area of maternal and child health. However, despite the promise, evidence for impact is severely lacking.[11][12] One 2011 systematic review by Noordam and colleagues states that "mHealth presents a new and pervasive platform for addressing prenatal and newborn health," but also points out that a "relative scarcity of articles with a quantitative design challenged the ability to statistically corroborate the impact of mHealth".[2] Another 2011 systematic review by Tamrat and Kachnowski concluded that, "Robust studies providing evidence on the impact of introducing mobile phones to improve the quality or increase the use of maternal health services are lacking".[1]

Those studies that have attempted to evaluate the impact of mHealth interventions on maternal and child health have been either inconclusive, of a very small scale, or of a quasi-experimental nature. The remainder of this section details the existing published literature on quantitative evaluations of mHealth interventions for maternal and newborn health.

Both of the 2011 systematic reviews describe a large-scale study in Zanzibar, which has since been more formally evaluated.[13][14] Though the 2009 protocol planned five evaluation areas, the 2012 publication only discussed skilled delivery attendance, for which it found a highly significant OR of 5.73. Both 2011 reviews also discussed a large-scale study in Ghana conducted in

partnership with the Grameen Foundation. The Grameen Foundation's 2012 report states, "A primary goal was to conduct rigorous quantitative research to measure the impact of the Mobile Technology for Community Health (MOTECH) system on the health outcomes of pregnant women and newborns".[15] However, project delays and logistical problems prompted three fundamentally different randomisation and evaluation designs to be successively attempted, none of which, according to the same report, were successfully implemented in time to have a study group of sufficient size to accurately study the population with appropriate statistical power. As such, the programme's impact is currently unknown.

Tamrat and Kachnowski present several more cases, none of which possess sufficient rigour to confirm or reject the effectiveness of SMS interventions on maternal and child health. For instance, a large programme in Serbia deployed an automated SMS-based intervention, which delivered prenatal health support based on pregnancy stage. As of 2007, 3200 women were enrolled; however, no outcomes research has been published to date.[2] In addition, Tamrat and Kachnowski mention two smaller trials: first, an SMS-based reminder system for prenatal care visits on the Thai/Myanmar border, and second, an SMS-based prenatal health support trial in Bangkok. The Thai/Myanmar border study is a pre-post comparison of aggregate data in a single catchment area, which precludes clear causal inference. Meanwhile, the study in Bangkok enrolled 68 participants and did not find a significant impact of SMS on gestational age at delivery, fetal birth weight, preterm delivery or route of delivery (vaginal vs caesarean section). However, because of the small sample size, the researchers stated, "Further studies into pregnancy outcomes in larger groups of pregnant women should be done".[16]

Two additional interventions, which were not covered by the aforementioned systematic reviews, have also emerged in the literature. First, launched in February of 2010, Text4baby is a national programme in the USA,[17] which targets high-risk, low-income and minority women, and has reached over 400 000 people as of February 2013.[18] There are 114 messages in the pregnancy protocol and these messages provide information on a variety of topics such as prenatal care, influenza, immunisation, developmental milestones, breast feeding and others.[18] Despite the hundreds of thousands of enrollees in Text4baby, the only evaluation with a control group is described by Evans and colleagues. The study is a pilot evaluation of women initially presenting for care at the Fairfax County, Virginia Health Department. A total of 90 women completed both the baseline and follow-up surveys. Comparing changes from baseline to follow-up between the two study arms, the researchers found that women in the intervention arm were more likely to agree with the statement, 'I am prepared to be a new mother', but showed no significant differences on actual health knowledge.[19] Second, in a

2011 study in England, Naughton et al[20] conducted a randomised controlled trial of 207 pregnant women who were currently smokers and found that participants in the intervention arm were significantly more likely to set a quit date and have determination to quit. However, the increased levels in self-reported abstinence and continued validated abstinence were non-significant.[20] The authors concluded that a larger efficacy trial was warranted.

In sum, there is limited quantitative research on the effect of SMS interventions on maternal and child health outcomes. Evidence seems promising that such interventions can help mothers feel more prepared, though evidence on actual health outcomes is limited and unclear, and larger scale evaluations are needed. One of the primary barriers to conclusive evidence in this area is that health outcome data are generally lacking, and investigations tend to be inconclusive due to underpowering. It is thus currently unproven whether educational text messages can improve maternal and child health outcomes. Much better powered trials are needed to investigate whether the often-positive process outcomes lead to significant changes in health outcomes.

This study aims to fill this important gap in the literature, and has been structured in the following manner to ensure scientific rigour and policy relevance. First, the intervention is structured as a quasi-randomised controlled trial (qRCT) to isolate causal inference. Second, it is the largest qRCT to date in the world to evaluate the impact of SMS on final maternal and child health outcomes. Third, four-group treatment assignment is conducted to comparatively evaluate different groups of SMS messages, that is, good household prenatal practice (GHPP) messages, care seeking (CS) messages, the combination of both types of messaging, and the control group. Fourth, measures of psychological and behavioural changes are designed to understand the mechanisms by which SMS can impact health outcomes. Fifth, the study incorporates advice from the literature about optimal SMS implementation by tailoring their timing, facilitating two-way contact through a hotline and sending the messages through medical providers.

METHODS

Figure 1 below illustrates the overall study design. On agreement with the Xi'an Health Bureau in Shanxi Province, China, Gaoling and Lantian were selected as two pilot counties. In each county, the local maternal and child health centre (MCHC) was eligible to be a study site. Public health professionals in each MCHC were eligible to recruit the participants after training. There were two inclusion criteria: local pregnant women must (1) own a cell phone in the household, and (2) visit a MCHC for antenatal care (ANA) during pregnancy. Women were enrolled during any visit to the ANC at the MCHC at any gestational age. The majority of pregnant women were eligible as the coverage of mobile phones in the community was 98% in one site and 85% in another site.

At enrolment, prior to treatment assignment, a baseline survey was conducted. Next, a quasi-randomised factorial assignment placed each participant into one of four possible message package programmes.

We chose factorial design for quasi-randomisation (qRCT) and the four study arms include: (1) GHPP messages, including advice on nutrition, exercise, self-awareness of depression, breast feeding, etc; (2) CS messages, which include information about government-

Figure 1 Flow chart of the newborn health project. ANC, antenatal care; CS, care seeking; GHPP, good household prenatal practice.
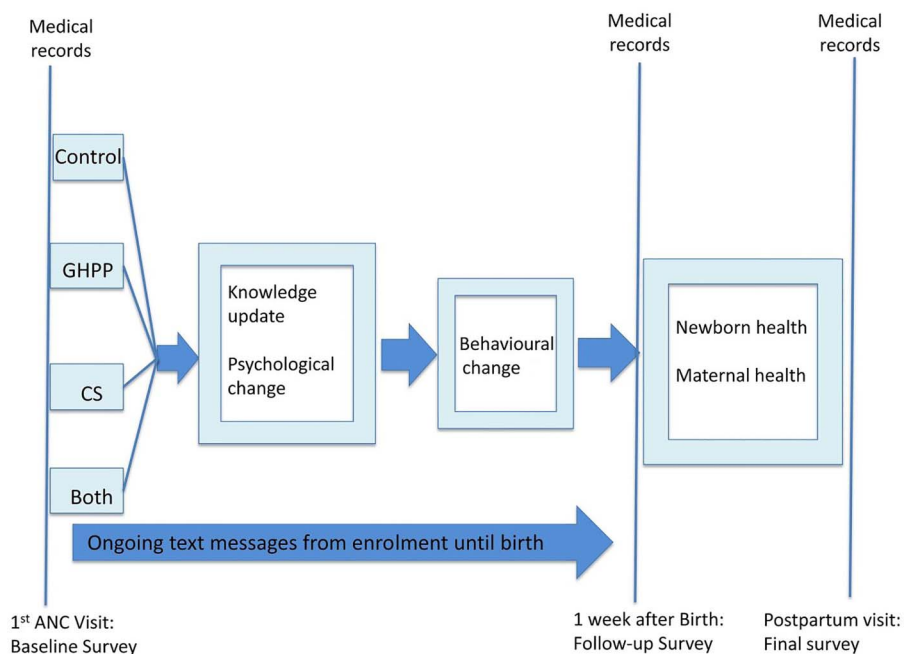
subsidised programmes, warning signs of potential problems and the importance of CS during illness; (3) both types of messaging; and (4) a limited set of 'status-quo' messages about pregnancy, which acts as a control. These control messages include updates on fetal development in different gestational stages, reminders for prenatal visits and promotion of certified skilled attendance during labour. For comparability, all groups receive these 'status-quo' messages. The factorial design enabled us to obtain evidence about the effects of SMS messages from far fewer participants than would be needed if GHPP and CS were individually tested in separate trials.[21]

All participates in the study have signed the informed consent documentation (see online supplementary additional file 1). Health workers at the MCHC obtain written informed consent from pregnant women who visit the MCHC. Any changes in our protocol implementation, including study site, sample size and timeline will be updated in the trial registry and reported to our sponsors and supporters.

The text messages are sent from the time of enrolment until delivery, and the contents are tailored according to the women's gestational week. A week after delivery, a follow-up survey is conducted, measuring knowledge, psychological and behavioural changes, as well as other pregnancy-related questions. Finally, a final survey is conducted a month after delivery to assess postpartum depression. Besides surveys, data from medical records are collected at baseline, during pregnancy, at birth, as well as 1 month after birth.

Totally, 148 messages have been designed in this study (table 1). Table 2 shows example messages from each category. It should be noted that though presented in English here, the messages that women receive are actually in Mandarin. According to our baseline survey, more than 99.8% pregnant women from the pilot counties received education from primary school and above. Therefore, all the participants of this study were capable of reading the Mandarin messages.

In the CS group and the group receiving the entire bank of messages, the reminders for prenatal visits and hospital delivery are more sophisticated. Additionally, more messages are sent to encourage the uptake of ANC than in the control or GHPP groups. In sum, the control and GHPP groups include six reminders with brief information, while the CS and entire message groups include a similar six reminders with more detail, and two additional messages (table 3).

Text messages

The first version of the message bank was developed based on an education package donated by Apricot Forest, Inc, for academic use, in 2011. With additional literature review on maternal health education, we ultimately designed 200 messages in 2012. The topics ranged from prenatal lifestyles (ie, nutrition, prenatal smoking and drinking, etc), fetus developmental stage, neglected issues (ie, postpartum depression, pain management practice, etc), suboptimal practice (ie, caesarean delivery, breast feeding, etc) and medical signs for seeking clinical care.

Table 1 Short message service (SMS) messages, by randomised group and time of delivery

Randomised group	Message categories	Sign-up day	First trimester	Second trimester	Third trimester	Final day
		Message delivery time and total number of SMS messages				
Control group (25)	Fetal development (19)	2	6	6	3	2
	Reminders for prenatal visit and hospital delivery (6)		1	1	4	
Care-seeking group (82)	Fetal development (19)	2	6	6	3	2
	Reminders for prenatal visit and hospital delivery (8)		2	2	4	
	Recognition of danger signs (45)		5	23	17	
	Reminders for government-subsidised projects (10)		3	2	5	
Good household prenatal practice group (91)	Fetal development (19)	2	6	6	3	2
	Reminders for prenatal visit and hospital delivery (6)		1	1	4	
	Healthy lifestyle (nutrition, physical activity, etc; 37)		15	16	6	
	Mental health during pregnancy (8)		1	4	3	
	Pain management (9)			4	5	
	Labour (6)			3	3	
	Breast feeding (6)				6	
Full SMS bank (148)	Full bank (148)	2	32	60	52	2

Table 2 Examples of SMS messages delivered to each study group

Group	Type	Time	Messages
Control	Prenatal visit reminders	Week 7, day 4	Regular antenatal care visits are the best way to protect both you and your baby. For normal pregnancy, experts recommend 1 antenatal visit per month in the first 7 months, 1 visit every 2 weeks between eighth and ninth months, and weekly visits in the last month of pregnancy.
Care-seeking	Reminders of government programmes	Week 7, day 2	Please pick up your free prenatal health management card from the county MCHC or township hospital within 3 months after pregnancy. With this card, you will receive 5 free prenatal examinations in the county MCHC during your pregnancy.
	Recognition of danger signs	Week 9, day 7	During pregnancy, if you ever experience high fever, dizziness, headache, vomiting (except morning sickness), blurred vision, vaginal bleeding, abdominal pain, rupture, fetal abnormalities, etc, please go to see a doctor.
Good household prenatal practice	Healthy lifestyle	Week 11, day 2	Calcium to promote bone development comes from the mother, so you may need calcium supplementation throughout your whole pregnancy. Milk, shrimp and soy products (if not allergic) are good sources of calcium. Please take advice if doctors suggest you supplement calcium, but be careful to avoid overdose.
		Week 10, day 6	Alcohol intake and smoking may cause birth defects, low intelligence and behavioural or learning disabilities. Please avoid smoking and any alcoholic drinks during pregnancy.
	Mental health	Week 6, day 6	Do you feel moody or worried? This may be due to the hormonal changes in your body. Take a deep breath, smile, picture in your mind what the baby would look like and talk with friends. Participate in activities that can help reduce stress.
	Non-medical pain management	Week 15, day 6	Breathing training can help you stay relaxed and calm during childbirth. Lamaze breathing exercise is the most commonly used way to reduce pain during labour. It's not so complicated! Just take a deep nasal breath with your nose, and then exhale slowly through your mouth. Please consult the maternal and infant healthcare centre for details.
	Labour knowledge	Week 30, day 7	Do you have confidence in delivery? Do you feel a little worried or panicked? Share your concerns with your doctor, and learn more information about labour in advance.
		Week 33, day 7	Many mothers said that the moment of childbirth feels like a bowel movement. During labour, if a doctor told you to push, you could just make effort as you do in a bowel movement.
	Breast feeding	Week 39, day 7	Breast feeding will be beneficial for you and your baby's health. The WHO recommends exclusive breast feeding for 6 months after birth, and continuing breast feeding for up to 2 years. Complementary feeding should start from 6 months of age.

MCHC, maternal and child health centre; SMS, short message service.

The next step was the localisation of messages for the two study sites: Gaoling and Lantian Counties, Shanxi Province, China, from March to July 2013. We conducted cognitive interviews of the first version of text messages among 33 new mothers and 15 health workers in Gaoling and 32 new mothers and 9 health workers in Lantian. Based on their feedback, we tailored the messages according to their local conditions and understanding of terminology. We also held a discussion session, in which we invited professors in maternal and

child health, local leaders from the MCHC, and government officials, to finalise the bank of text messages.

The final version of text messages includes 148 messages: 34 messages in the first trimester (sign up date and gestational weeks 5–12), 60 messages in the second trimester (gestational weeks 13–27) and 54 messages in the last trimester (gestational weeks 28–40; table 1). It is possible that the differences in outcomes among randomised groups are partially attributed to the number of messages. We tried to create the 'placebo' by sending

Table 3 Reminders for prenatal visits and hospital delivery delivered to each study group

Week	Control and GHPP groups	CS and full message groups
Week 7, day 4		Regular antenatal care visits are the best way to protect both you and your baby. For normal pregnancy, experts recommend 1 antenatal visit per month in the first 7 months, 1 visit every 2 weeks between eighth and ninth months, and weekly visits in the last month of pregnancy.
Week 10, day 2	From this week, you can start your first antenatal visit and establish your own health profile. Routine examination includes urine examination, blood type, liver and kidney function, blood sugar, hepatitis B test, etc.	From this week, you can start your first antenatal visit and establish your own health profile **for free**. Routine examination includes urine examination, blood type, liver and kidney function, blood sugar, hepatitis B test, etc. Don't forget to go to the hospital!
Week 16, day 6	The next 2 weeks you could go for the second prenatal visit. This visit should include screening for Down syndrome, blood pressure and other neural tube defects.	The next 2 weeks you need to get the second prenatal visit **for free**. This visit should include screening for Down syndrome, blood pressure and other neural tube defects.
Week 17, day 2		Your doctor may advise you to have ultrasound diagnosis this week. Several ultrasound diagnoses are needed during pregnancy. It can help to check your baby's organ development, determine fetal position and placental location, and assess placental function, etc.
Week 28, day 2	This week, you could go for the third prenatal visit. This visit should include general physical examination, blood pressure, etc.	Starting from this week to the 36th week, you can have an antenatal visit once every 2 weeks. After 36 weeks you can have it once a week.
Week 34, day 2	Now you could go for the 4th prenatal visit. This visit may include ultrasound diagnosis, general physical examination, blood pressure, etc.	You can have ultrasound diagnosis during the next 2 weeks to observe the growth and development of your baby, the location and maturity of placenta and amount of amniotic fluid. In addition, it can capture if there is organ abnormality during late pregnancy or if it was missed in the second trimester of pregnancy.
Week 36, day 2	Now you could go for the fifth prenatal visit. This visit may include ultrasound diagnosis, general physical examination, blood pressure, etc.	Starting from this week, you can have a weekly antenatal visit to ensure the safety of you and your baby. It is the 'final stage' of your pregnancy, go for it!
Week 38, day 4	From now on, please be prepared to go to the health facility (county maternal and child health centre or country hospital) for delivery.	Doctors will decide whether you should have early admission, fetal monitoring and your delivery approach by measuring the pelvis and ultrasonography. If doctors suggest that you should do vaginal delivery, please don't say no just because of fear of pain.

CS, care seeking; GHPP, good household prenatal practice.

messages to the control group and we admit the difference in the number of messages is one of our limitations. We reviewed the literature about the effect of total number or frequency of text messages and discuss the limitation in future publications. Finally, an algorithm was developed for automatic delivery of text messages by randomised group and by gestational date.

Study site

Our project office had a strategic planning meeting with Xi'an Development and Reform Commission (DRC) in 2012 in order to select sites. We selected two counties, Gaoling and Lantian, to conduct the large-scale experiment in Shanxi Province, China. According to County Government statistics in 2013, Gaoling County has a population of 330 000, with 3441 pregnant women in 2012. In addition, 98% of the households in Gaoling own a cell phone. Lantian County has a population of

640 000, with 5901 pregnant women in 2012. The rate of cell phone ownership in Lantian is 85%. The study in Gaoling and Lantian Counties has been incorporated into the government agenda within Xi'an DRC to improve maternal and child health. With support from the Health Bureau, we mainly partner with MCHC for recruitment of pregnant women, sending text messages, and surveying the participants to measure outcomes. Meanwhile, we also partner with county hospitals to collect medical data for women who deliver in a county hospital. The MCHC collects, assesses and manages solicited and spontaneously reported adverse events and other unintended effects of trial interventions or trial conduct and reports this information to the Newborn Health Project Office in a timely manner. The sponsor conducts audits and evaluations of trial conduct every 6 months. After generating the scientific results from the pilot, we will host a meeting with the provincial

government to discuss the scale-up of this pilot to Shaanxi Province. For a visual depiction of study operations and partnerships, see figure 2 below.

Pilot study

Between July and August 2013, we trained 20 local public health professionals and 4 student researchers regarding the consent process, cognitive debriefing, face-to-face interviewing and phone interviewing. Pretesting of the information technology (IT) system to deliver messages began in June 2013. We collected 10 staff members' birth dates, telephone numbers, and hypothesised gestational weeks to input into the IT system, and then we tested the quasi-randomisation process and message delivery. Between 29 August and 30 September 2013, we began to test the designed survey questionnaires, including baseline survey and two follow-up surveys. In September and October 2013, we conducted pilot testing among 140 women with the aim of optimising recruitment, treatment assignment, message sending, survey procedures and measurement accuracy. Those women were included in our test and they will be excluded in programme evaluation. The survey questionnaires were finalised after incorporating feedback from the testing.

Treatment assignment

Treatment assignment was performed by an IT platform. First, once the pregnant women were enrolled, demographic characteristics were entered as a part of pregnancy records and all participants were randomised into one of the four groups according to the month and day of birth (table 4). The algorithm assigned women to one of four treatment groups based on whether their month and day of birth were both even, both odd, month even and day odd, or month odd and day even.

Both health workers enrolled the participants, and the participants were blinded to the assignment method. Month and day of birth of the participants were entered by health workers into the computer and all of the participants were allocated into four groups by computer. The health workers were not informed either about the assignment method or the allocation outcomes. The participants were not informed about the allocation either. In this qRCT, adequate concealment is ensured by having the allocation of participants to different treatment groups conducted by a computer algorithm. This trial is qRCT but double blind in treatment assignment. Unblinding at the individual level is not permitted during the trial. The trial did not involve stratification or blocking.

Outcome measures
Primary health outcome

The primary outcome is newborn health measured by the appropriateness of weight for gestational age. We define this as being neither small for gestational age (SGA) nor having macrosomia (weighing ≥4000 g at birth). Data are from both medical records at birth and self-reports by mothers in the endline survey. The outcome measure was informed in the consent form and was not blinding. We hypothesised that SMS might reduce inappropriate weight for gestational age, including SGA and macrosomia. For example, we sent messages with nutrition advice during their pregnancy, which might reduce SGA. Meanwhile, because rural

* EMI: Evaluation for mHealth Interventions
** MCHC: Maternal and Child Health Centre
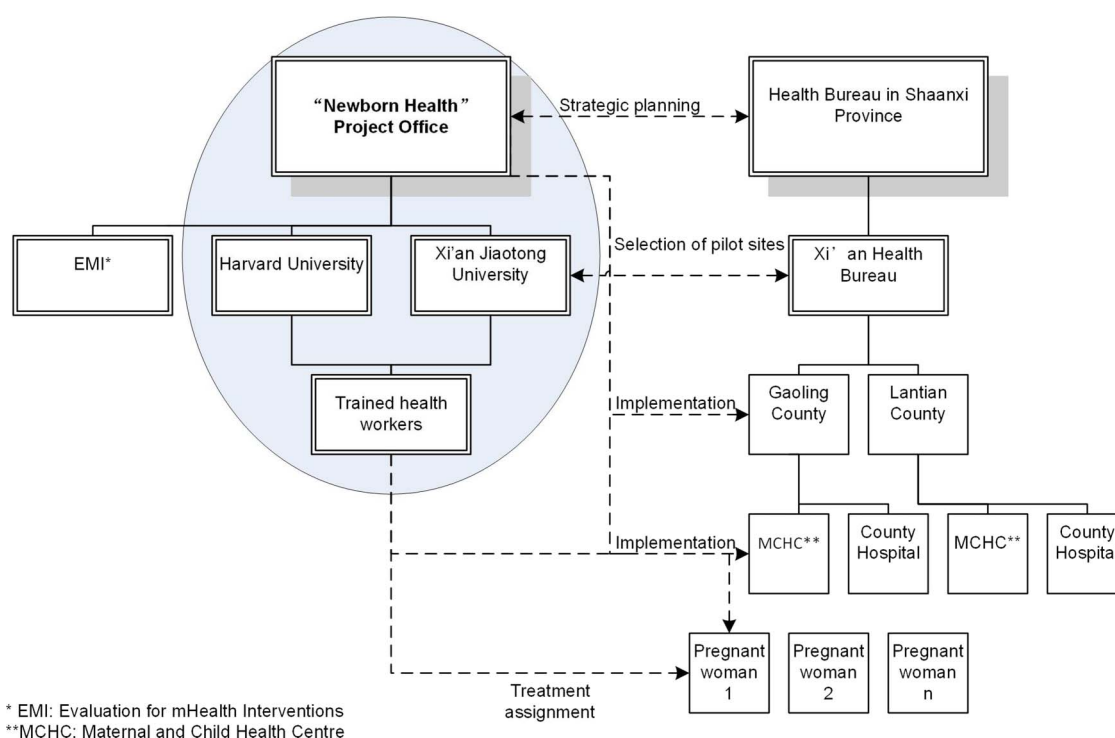
Figure 2 Partners for the study.

Table 4 Study group assignment

Randomised group	The month of birth	The day of birth
Control	Odd	Odd
GHPP	Even	Odd
CS	Odd	Even
GHPP+CS	Even	Even
CS, care seeking; GHPP, good household prenatal practice.		

China households have a preference for large babies, we delivered info about appropriate fetal development, which might reduce macrosomia.

Secondary health outcomes

There are important secondary outcomes to reflect maternal and newborn health. First, maternal health will be evaluated by measuring changes in perceptions of general health and postpartum depression. The level of severe maternal morbidity will be measured during pregnancy and childbirth through the summary indicator, 'near-miss', including information such as cardiovascular, respiratory or renal dysfunction. Rather than the WHO recommended 'near-miss' indicator, we will be using the one created by Reichenheim and colleagues in 2009. In their 2009 systematic review of near-miss literature, Reichenheim and colleagues created a set of 13 indicators derived from the indicators most commonly used in the 51 source studies reviewed. The authors intended this set to be universally usable.[22] Though ideally universal, this set is perhaps less specific than the WHO criteria, and was measured to categorise six times as many cases as 'near-miss' as the WHO criteria in a Brazilian study.[23] However, while we are able to collect information on the Reichenheim indicator in county MCHCs and hospitals on the study sites, the information required by the WHO indicator includes highly technical data only partially available in rural China.

Second, the level of severe neonatal morbidity will be measured through a summary indicator, the neonatal adverse outcome indicator (NAOI). The NAOI is a measure proposed by Lain et al[24] in 2012, and was found to be a better predictor of death and readmission than a 5 min Apgar <7. The NAOI includes 15 diagnosis-based indicators (such as death, gestational age <32 weeks, respiratory distress syndrome, seizure, etc) and 5 procedure-based indicators (resuscitation, ventilatory support, central venous or arterial catheterisation, blood transfusion, and pneumothorax intercostal catheterisation).[24]

Psychological and behavioural outcomes

Process indicators mainly focusing on psychological and behavioural outcomes were also measured. First, the psychological outcome will include nine dimensions: attitudes, personal norms, self-efficacy, social desirability, intentions, plans, susceptibility, expectations and severity. Second, behavioural outcomes including the actual number of prenatal visits over expected visits, and the uptake of government-subsidised programmes (eg, duration of folic acid and uptake of infant vaccinations), nutrition, moderate exercise, CS when ill and caesarean section were measured during the course of the pregnancy and childbirth.

Sample size

For the primary outcome indicator, 'appropriate weight for gestational age', we assumed that the Chinese national average for SGA estimated by Lee et al,[25] 6.5%, and the Chinese national average rate of macrosomia estimated by Lu et al,[26] 7.83%, would sum to the expected rate of inappropriate weight (14.33%) in the control group. Between the four groups, we plan to make four of the six possible two-way comparisons: the GHPP group to control, CS to control, GHPP to CS and the group with both messages to whichever message group comes out best: GHPP or CS. The required sample size was based on powering the control to either GHPP or CS treatment comparisons while using a Bonferroni correction for four hypothesis tests. Bonferroni correction is used to counteract the problems that typically arise with multiple comparisons.

For the actual analysis, we intend to use the Bonferroni-Holm correction, but for prospective power calculations, we used the regular Bonferroni correction. This is because even though it weakens power, it is always at least as stringent as the Bonferroni-Holm and can be incorporated into prospective power calculations while retaining a closed form solution. With this correction, we found that a sample size of 5200 is required to identify a reduction of 4.33 percentage points (down to 10%), with 81% power, α of 0.05 and using a two-sided test. R code is available on request for power calculations.

The interpretation of our power calculations is that our experiment is powered to detect four differences of 4.33% under a Bonferroni correction, though it seems unlikely that all differences will be so large. It is important to note that the incremental nature of the Bonferroni-Holm algorithm (as opposed to a pure Bonferroni correction) to some extent accounts for this. However, the stepwise nature of the procedure makes power along subsequent steps difficult to prespecify.

Data collection, storage and access

The study data are mainly collected in three rounds of surveys and from medical records. The baseline survey is administered by a health worker at the local MCHC in the first antenatal visit, after obtaining the written informed consent from women who agreed to participate in the study. The follow-up surveys will be conducted at home at 1 week after delivery by a health worker. The final survey is a phone interview at 1 month after delivery. In addition, the clinical data will be collected from local MCHC and county hospitals. The IT platform developed in this study has the capacity for collecting medical records. Four modules were designed to

collect medical records: antenatal check-ups, postnatal check-up, high-risk pregnancy and other tests (ie, blood test, ultrasound).

Although all women were enrolled at the local MCHC, some women may choose to deliver in a local county hospital. For women who deliver in facilities other than MCHC and county hospitals, an additional survey will be conducted to complement the information from medical records. For each participant, the survey data, medical record data and the treatment assignment were matched by telephone number, and if matching was unsuccessful, by the combination of name and village.

The study paper questionnaires are stored at the project office in Xi'an Jiaotong University. Random checks of sampled data entry were applied to guarantee the data input quality. Original questionnaire checking will be conducted if significant issues are found during the data cleaning process. The personal information of study participants will only be accessible to the data management staff; other study implementation team members, including data analysts, do not have access to it. No formal data monitoring committee was established, and the project co-principal investigator (ZZ) has supervised data quality and has access to the final trial data.

Data analysis strategy

All expectant women were randomly assigned into four groups: (1) both GHPP and CS; (2) only GHPP without CS; (3) only CS without GHPP and (4) control. Accordingly, we propose a four-group balance check to enable us to compare among four subgroups, as seen in table 5.

We will conduct a balance check of preintervention observable characteristics that indicates whether the procedure of quasi-randomisation successfully balances among the study arms.

It is possible that the enrolled women were the relatively better off group, compared with the counterpart who had no cell phone at baseline. We will look at both population and sample distributions for demographic characteristics at baseline to speak to the representativeness of the sample.

To analyse the effectiveness of the different treatment arms with regard to the primary outcome, multivariable logistic regression using an intent-to-treat (ITT) design will primarily be applied. Accordingly, four types of impact could be estimated:

$$Yi = \beta0 + \beta1 GHPP + \beta2 CS + \beta3 BOTH + Xi + Ci + \varepsilon i$$

Table 5 Study group variable scheme

		GHPP	
		1	0
CS	1	(CS, GHPP)	(CS, no GHPP)
	0	(no CS, GHPP)	(no CS, no GHPP)

CS, care seeking; GHPP, good household prenatal practice.

Yi indicates the dependent variables, and i represents the individual participant. GHPP, CS and BOTH are dummy variables that take the value 1 if a person got that version of messages and 0 otherwise. The reference group is the control group. Xi is the demographic controls and Ci is county fixed effects.

We choose to estimate β3BOTH as an independent third group rather than estimating the main effects of GHPP and CS while estimating β3GHPP×CS as an interaction effect on the assumption that the combined effect of the two SMS interventions will be subadditive. Estimating the equation in this fashion allows us to then test directly whether the effect of receiving both is statistically greater than receiving either separately by a simple t-test of the difference between β3 and either β1 or β2, which is of more direct interest to the investigators than testing whether the effects are linearly additive.

From this regression, four hypotheses will be tested, and the Bonferroni-Holm algorithm will be used to correct for implementing four tests. First, whether either message set alone had a significant effect will be captured by β1 and β2 for GHPP and CS, respectively. Next, whether either GHPP or CS alone is superior to the other is captured by a t-test to determine whether β1−β2 is significantly different from zero. Finally, we will test whether receiving both interventions together is superior to only receiving whichever of GHPP or CS is estimated to have the larger effect. Though which of the two groups is chosen for comparison will be data driven, making the other comparison will either be redundant by transitivity after β1 and β2 are compared with each other or will be left undetermined by the investigators, and in either case will not warrant an independent contribution to the Bonferroni-Holm correction process.

We also considered two other alternate correction procedures besides the Bonferroni/Bonferroni-Holm procedure which would have more explicitly accounted for these co-dependencies arising from our factorial design. The first was Tukey's honest significant difference test. The issues with this test are that (1) it assumes normally distributed outcomes, whereas our outcome of interest is dichotomous, and (2) it makes all possible group comparisons, whereas given our factorial design, four would be of sufficient interest. The other was the student-Newman-Keuls test, which would be ideal in construction except that (1) it does not necessarily control the family-wise error rate and (2) cannot be used in a priori power calculations due to its sequential nature and the resulting ambiguity in its probability of making a type I error.

Analysis of the secondary outcomes will vary by outcome. Analysis of neonatal morbidity will also proceed by analogous logistic regression. Maternal near-miss is a rare event for which large sample Wald estimates of CIs are inappropriate, so in that case exact CIs and p values for effect size will be determined via simulation from the binomial distribution. The effect of the messages on the categorical, psychological and behavioural measures will be estimated with logistic or

ordered logistic regression, as appropriate; these will then inform subsequent structural equation modelling for the mediation analysis of text messages on behaviour. Potential linkages from behaviours to health outcomes will be analogously explored.

The rationale for ITTs is to estimate the effects of intervention policy rather than the effects among the complying participants. Often ITT estimates are useful from a policy standpoint and non-adherence does not disallow ITT analyses. However, non-adherence does differentiate the resulting inference of ITT from treatment on the treated (ToT) analyses. In this case specifically, the inference of our results will be what happens if a party (be it hospital, government or other actor) signs up a set of women to receive a set of text messages (ITT), rather than what happens if a set of women both receive and read a certain set of text messages (ToT). With no non-adherence, the two would be identical, but ITTs are better than ToTs specifically and exactly because they make non-adherence a non-issue. It is the former rather than latter that is realistically in the power of the relevant future implementers of our protocol.

We will also conduct ToT analysis given that we collected data about participants who requested the text messages to be discontinued in treatment phase and who reported that they received the texts and/or read them in follow-up surveys.

Both ITTs and ToTs might suffer from missing outcomes. To address missing outcomes, we can either impute the missing outcomes through listwise deletion or multiple imputation. We can use one strategy as our main and the other as our sensitivity check. We can also focus on the subset of participants whose pretreatment characteristics are non-significantly different and who have been observed for outcome measures.

Our quasi-randomisation leaves us open to the possibility, however unlikely, that any random shocks at the level of the mother's birth month or birth day that would later affect the mother's newborn though unobserved covariates will become correlated with treatment assignment. Therefore, as a sensitivity check, we will re-estimate our main analysis in a jack-knife-like fashion, dropping all women born in each month successively to check if a month effect can account for our main results.

Further, we will document the facts regarding implementation challenges based on the data we collected in three rounds of surveys. For instance, we collected data regarding treatment dropoff, SMS reception, reading and peer communication to understand the challenges in SMS projects.

DISCUSSION
Lessons learnt
A number of studies have investigated how SMS programmes can be designed for optimal effectiveness, and several lessons have emerged.

Numerous studies have found, either via surveys, interviews, focus groups or experience, that personal tailoring seems to be preferred by message recipients.[15 20 27–29] Tailoring can involve including the recipient's name, their child's name, personalised information about goals or health recommendations, information tailored to gestational age, information on their particular health provider, or other tailored content. Recipients also seem to prefer that the messages be relatively simple and concise.[20 30 31] For example, in a survey of 190 parents regarding immunisation reminders, the majority of parents preferred reminder messages that relayed which vaccination was due, on what date and their provider's phone number over messages that contained none of these elements and over messages that contained all these elements plus additional information on the DTaP vaccine.[27]

SMS programmes can be unidirectional or involve two-way communication. Some authors have found that two-way systems may be preferable. One study found that two-way communication was important to maintaining a 'human aspect' to the intervention.[28] Another study suggested that, "inviting a reply would make 'you think more about the text message' otherwise 'you don't have to do anything with it so you read it and then forget about it'".[20] In Gurman et al's[32] 2012 review, 12 of the 16 included studies had a two-way SMS design, and in Fjeldsoe et al's 2009 review, all 14 included trials had a two-way component. The latter study notes that there were no clear differences in intervention outcome based on whether participants or the researcher initiated dialogue, but that all the preventive health behaviour studies used researcher-initiated techniques.[10]

One potentially important and related finding comes from a focus group study on influenza vaccination in New York City. The study found that women's most trusted source of health information is their medical provider. Others preferred to get information from the internet, news media, family and friends. Most women reported that messages regarding vaccine safety or benefits would not directly change their beliefs, but would encourage them to discuss the influenza vaccine with their provider.[33] This may imply that any given SMS intervention might work best when participants have access to a medical provider with whom they can discuss the messages and who will corroborate their contents.

To summarise the lessons learnt, text messages in our study have been tailored by gestational age to the individual recipient. Two-way communication may be beneficial, and reinforcement, or confirmation, of messages from healthcare providers through hotline may play a key role. In our study, we intend to heed to these lessons for maximum impact. For instance, the hotline already existed before we launched the programme, which was staffed 8 hours a day for 7 days a week by health workers at MCHCs. We provided the hotline phone number of the MCHC through SMS during the trial to encourage patient-initiated contact with healthcare providers.

Participants can use the hotline service to seek relevant prenatal care and interventions during the trial. This feature is included to ensure that participants are not prohibited from receiving needed healthcare throughout the trial. The message sending was terminated if the participant sent a 'stop' message to our platform.

Potential contributions

Beyond the investigation of health outcomes, the study will also explore via survey what emotional and mental processes inspire women to act on or ignore the health messages they receive. Thus, the study will provide insight into what types of messaging and framing can affect behaviour the most. This kind of evidence could be highly valuable to anyone involved in the burgeoning field of mHealth delivery; and more specifically, the project will extend the current body of knowledge among scholars and policymakers in the area of impact evaluation for maternal and newborn health interventions.

Furthermore, this study has broader implications amidst sweeping health reforms in China. Given recent emphasis by the Central Government on preventive care in rural areas, this SMS-based intervention, if proven effective, could have widespread appeal given its advantages over current methods to reach women during pregnancy. Traditional maternal education consists of posters, fliers and brochures,[2] which might provide quality and trust but falter in the areas of cost, timing, access and digestibility.

From a policy perspective, it is useful to understand which components of the bank of SMS messages should be scaled up. Therefore, it is important to disentangle which component contributes most to final neonatal health outcomes. Taken together, are all the components of the SMS bank effective in changing maternal behaviour and enhancing neonatal health? Which mechanisms are at play: GHPP, CS in pregnancy or both? This study may provide evidence for an SMS-based approach, which addresses the weaknesses of the status quo. If proven effective, this intervention could be integrated into maternal and child primary care delivery and scaled to low-resource counties across China and beyond.

Several challenges have been witnessed over the course of implementation and early data collection stages of this study.

First, local conditions can make study implementation difficult. For example, in the early stages of this study, there have been more challenges to implementation in the lower income county, Lantian County, where an electronic platform was not available until January 2014. This is in contrast to Gaoling County, where access to an electronic platform was available immediately.

Second, there has been limited capacity of the local health workers in evaluating the effectiveness of the intervention. In China, the MCHC plays an important role in managing and educating pregnant women. Without the MCHC staff, recruitment of pregnant women at a large scale in this study would not be possible. However, as the early stages of this study have revealed, MCHC health workers cannot be solely tasked with collecting data for evaluation due to their heavy workload and tight daily schedule. For optimal implementation of this kind of intervention and evaluation, a hybrid system, merging researchers and health workers, is vital.

Third, there is the challenge of losing participants to follow-up or simply a refusal to participate due to cultural reasons. Thus far, this study has seen a small proportion (3%) of local women refuse to participate in the study. Since the intervention lasts for the whole pregnancy, the attrition over time might be a significant concern.

Fourth, one of the most important unexpected challenges was that many women in the study delivered their newborn outside of the MCHC network. According to the summary statistics from MCHC in 2012, even though 90% of rural women went to MCHC for prenatal visits, only 50% of pregnant women delivered at a MCHC. The study would be significantly influenced if the high-risk pregnancies shifted to other health facilities such as the county hospital or tertiary hospitals for childbirth. In terms of programme evaluation, this unexpected problem suggests potential selection bias in outcome measures. We have developed two approaches to manage this problem: First, we collaborated with the county hospital to follow-up with the new mothers enrolled in our programme. Around 40% of women delivered at the county hospital and around 10% of women went to other health facilities. Second, in the first home visit by health workers, which covers all new mothers, an additional survey was conducted for women who delivered outside a MCHC.

In sum, this pilot is the first large-scale effort to build a comprehensive evidence base on the impact of prenatal text messages via cell phone on maternal and newborn health outcomes in China. This study includes both subjective and objective health outcome measures. Beyond this, psychological and behavioural indicators are measured to further understand the underlying mechanisms in improving health outcomes. The study has broad implications for public health policy in China and the implementation of mHealth interventions in low-resource settings around the world.

Trial status

All participants in the study have signed the informed consent. The first version of the message bank was developed in February 2013. Pretesting of the SMS message bank was completed in July 2013. We offered training for health workers in July and August 2013. The study has enrolled participants since September 2013 and the intervention ended in October 2015 (see online supplementary additional file 2, estimated enrolment and treatment allocation). We have communicated protocol

modifications to investigators, the Institutional Review Board (IRB), trial registries and study sponsors.

Dissemination policy

We registered the study in the protocol registration system, the service of the US National Institutes of Health, which is recognised internationally. The full protocol is accessible by the public (see online supplementary additional file 3). We plan to communicate trial results to participants, healthcare professionals, policy-makers, the funder, the public and other relevant groups via conferences, publication or other data sharing arrangements.

Authorship eligibility guidelines

We follow the 'Guidelines for Investigators in Scientific Research' for authorship eligibility. We follow two critical safeguards to enhance accuracy and scientific rigour in publication: (1) requirement of active participation of each co-author in verifying the part of a manuscript that falls within his/her specialty area, and (2) the designation of one author who is responsible for the validity of the entire manuscript. Individual who has made a significant intellectual or practical contribution is eligible to be considered as co-authors. The first author should assure that he/she has reviewed all the primary data on which the report is based and provide a brief description of the role of each co-author. Appended to the final draft of the manuscript should be a signed statement from each co-author indicating that he/she has reviewed and approved the manuscript to the extent possible, given expertise. We do not accept 'honorary authorship'. We do not intend to use professional writers. More details can be found through http://hms.harvard.edu/about-hms/integrity-academic-medicine/hms-policy/faculty-policies-integrity-science/guidelines-investigators-scientific-research.

Author affiliations

[1]Freeman Spogli Institute for International Studies, Stanford University, Stanford, California, USA
[2]Nutrition Department, Harvard School of Public Health, Boston, Massachusetts, USA
[3]School of Public Policy and Administration, Xi'an Jiaotong University, Xi'an, Shaanxi, China
[4]Global Health and Population Department, Harvard School of Public Health, Boston, Massachusetts, USA
[5]Vera Solutions, Cambridge, Massachusetts, USA

Contributors YS has led conceiving and designing the intervention and study, led literature review and detailed the method of programme evaluation. She serves as the Project Director for the team. CY helped conceive the study, participated in literature review, and led the development of the text messages and survey questionnaires. ZZ helped design the study and has led the implementation of the project. JH conceived the psychological aspects of the study, helped conceive behavioural aspects of the study, participated in literature review, participated in development of survey questionnaires and led power calculations. BC participated in literature review and helped conceive behavioural aspects of the study. All the authors have made significant intellectual or practical contributions. All the authors drafted the manuscript with collaborative effort. All authors critically revised the manuscript and approved the final version of this paper.

REFERENCES

1. Noordam AC, Kuepper BM, Stekelenburg J, *et al*. Improvement of maternal health services through the use of mobile phones. *Trop Med Int Health* 2011;16:622–6.
2. Tamrat T, Kachnowski S. Special delivery: an analysis of mHealth in maternal and newborn health programs and their outcomes around the world. *Matern Child Health J* 2012;16:1092–101.
3. Free C, Knight R, Robertson S, *et al*. Smoking cessation support delivered via mobile phone text messaging (txt2stop): a single-blind, randomised trial. *Lancet* 2011;378:49–55.
4. Blake H. Innovation in practice: mobile phone technology in patient care. *Br J Community Nurs* 2008;13:160.
5. Furber GV, Crago AE, Meehan K, *et al*. How adolescents use SMS (short message service) to micro-coordinate contact with youth mental health outreach services. *J Adolesc Health* 2011;48:113–15.
6. Spielberg F, Kurth A, Reidy W, *et al*. Iterative evaluation in a mobile counseling and testing program to reach people of color at risk for HIV-new strategies improve program acceptability, effectiveness, and evaluation capabilities. *AIDS Educ Prev* 2011;23:110–16.
7. Guy R, Hocking J, Wand H, *et al*. How effective are short message service reminders at increasing clinic attendance? A meta-analysis and systematic review. *Health Serv Res* 2012;47:614–32.
8. Jacobson Vann JC, Szilagyi P. Patient reminder and recall systems to improve immunization rates. *Cochrane Database Syst Rev* 2005; (3):CD003941.
9. Cole-Lewis H, Kershaw T. Text messaging as a tool for behavior change in disease prevention and management. *Epidemiol Rev* 2010;32:56–69.
10. Fjeldsoe BS, Marshall AL, Miller YD. Behavior change interventions delivered by mobile telephone short-message service. *Am J Prev Med* 2009;36:165–73.
11. Wei J, Hollin I, Kachnowski S. A review of the use of mobile phone text messaging in clinical and healthy behaviour interventions. *J Telemed Telecare* 2011;17:41–8.
12. Chib A, van Velthoven MH, Car J. mHealth adoption in low-resource environments: a review of the use of mobile healthcare in developing countries. *J Health Commun* 2015;20:4–34.
13. Lund S, Hemed M. *Wired mothers: use of mobile phones to improve maternal and neonatal health in Zanzibar*. University of Copenhagen. Danish Research Network for Health (Enreca health), 2010.
14. Lund S, Hemed M, Nielsen BB, *et al*. Mobile phones as a health communication tool to improve skilled attendance at delivery in Zanzibar: a cluster-randomised controlled trial. *BJOG* 2012;119:1256–64.
15. Mobile technology for community health in Ghana: what it is and what Grameen Foundation has learned so far. 2012. http://www.grameenfoundation.org/sites/grameenfoundation.org/files/MOTECH-Lessons-Learned-Sept-2012.pdf (accessed 12 Sep 2015).
16. Jareethum R, Titapant V, Tienthai C, *et al*. Satisfaction of healthy pregnant women receiving short message service via mobile phone

for prenatal support: a randomized controlled trial. *Med J Med Assoc Thai* 2008;91:458.

17. van Velthoven MH, Majeed A, Car J. Text4baby–national scale up of an mHealth programme. Who benefits? *J R Soc Med* 2012;105:452–3.

18. Remick AP, Kendrick JS. Breaking new ground: the text4baby program. *Am J Health Promot* 2013;27:S4–6.

19. Evans WD, Wallace JL, Snider J. Pilot evaluation of the text4baby mobile health program. *BMC Public Health* 2012;12:1031.

20. Naughton F, Jamison J, Sutton S. Attitudes towards SMS text message smoking cessation support: a qualitative study of pregnant smokers. *Health Educ Res* 2013;28:911–22.

21. McAlister FA, Straus SE, Sackett DL, *et al.* Analysis and reporting of factorial trials: a systematic review. *JAMA* 2003;289:2545–53.

22. Reichenheim ME, Zylbersztajn F, Moraes CL, *et al.* Severe acute obstetric morbidity (near-miss): a review of the relative use of its diagnostic indicators. *Arch Gynecol Obstet* 2009;280:337–43.

23. Lobato G, Nakamura-Pereira M, Mendes-Silva W, *et al.* Comparing different diagnostic approaches to severe maternal morbidity and near-miss: a pilot study in a Brazilian tertiary hospital. *Eur J Obstet Gynecol Reprod Biol* 2013;67:24–8.

24. Lain SJ, Algert CS, Nassar N, *et al.* Incidence of severe adverse neonatal outcomes: use of a composite indicator in a population cohort. *Matern Child Health J* 2012;16:600–8.

25. Lee AC, Katz J, Blencowe H, *et al.* National and regional estimates of term and preterm babies born small for gestational age in 138 low-income and middle-income countries in 2010. *Lancet Glob Health* 2013;1:e26–36.

26. Lu Y, Zhang J, Lu X, *et al.* Secular trends of macrosomia in southeast China, 1994–2005. *BMC Public Health* 2011;11:818.

27. Ahlers-Schmidt CR, Hart T, Chesser A, *et al.* Content of text messaging immunization reminders: what low-income parents want to know. *Patient Educ Couns* 2011;85:119–21.

28. Fjeldsoe BS, Miller YD, O'Brien JL, *et al.* Iterative development of MobileMums: a physical activity intervention for women with young children. *Int J Behav Nutr Phys Act* 2012;9:151.

29. Kharbanda EO, Stockwell MS, Fox HW, *et al.* Text4Health: a qualitative evaluation of parental readiness for text message immunization reminders. *Am J Public Health* 2009;99:2176.

30. Militello LK, Kelly SA, Melnyk BM. Systematic review of text-messaging interventions to promote healthy behaviors in pediatric and adolescent populations: implications for clinical practice and research. *Worldviews Evid Based Nurs* 2012;9:66–77.

31. Whittaker R, Matoff-Stepp S, Meehan J, *et al.* Text4baby: development and implementation of a national text messaging health information service. *Am J Public Health* 2012;102:2207–13.

32. Gurman TA, Rubin SE, Roess AA. Effectiveness of mHealth behavior change communication interventions in developing countries: a systematic review of the literature. *J Health Commun* 2012;17(Suppl 1):82–104.

33. Kharbanda EO, Vargas CY, Castaño PM, *et al.* Exploring pregnant women's views on influenza vaccination and educational text messages. *Prev Med* 2011;52:75–7.