

Implementation and evaluation of nationwide scale-up of the Surgical Safety Checklist

M. C. White^{1,4,6} , K. Randall¹, N. F. E. Capo-Chichi², F. Sogas³, S. Quenum¹, K. Wright¹, K. L. Close¹, S. Russ⁵, N. Sevdalis⁵ and A. J. M. Leather⁴

¹Department of Medical Capacity Building, Mercy Ships Africa Bureau, ²Department of Paediatric Surgery, Centre National Hospitalier Universitaire Hubert Koutoukou Manga and ³Faculté des Sciences de la Santé de Cotonou, Université d'Abomey Calavi, Cotonou, Benin, and ⁴Centre for Global Health and Health Partnerships and ⁵Centre for Implementation Science, King's College London, and ⁶Department of Anaesthesia, Great Ormond Street Hospital for Children NHS Foundation Trust, London, UK

Correspondence to: Dr M. C. White, Centre for Global Health and Health Partnerships, Global Health Offices, Weston Education Centre, Cutcombe Road, London SE5 9RJ, UK (e-mail: doctormcw@gmail.com)

Background: The WHO Surgical Safety Checklist improves surgical outcomes, but evidence and theoretical frameworks for successful implementation in low-income countries remain lacking. Based on previous research in Madagascar, a nationwide checklist implementation in Benin was designed and evaluated longitudinally.

Methods: This study had a longitudinal embedded mixed-methods design. The well validated Consolidated Framework for Implementation Research (CFIR) was used to structure the approach and evaluate the implementation. Thirty-six hospitals received 3-day multidisciplinary training and 4-month follow-up. Seventeen hospitals were sampled purposively for evaluation at 12–18 months. The primary outcome was sustainability of checklist use at 12–18 months measured by questionnaire. Secondary outcomes were CFIR-derived implementation outcomes, measured using the WHO Behaviourally Anchored Rating Scale (WHOBARS), safety questionnaires and focus groups.

Results: At 12–18 months, 86.0 per cent of participants (86 of 100) reported checklist use compared with 31.1 per cent (169 of 543) before training and 88.8 per cent (158 of 178) at 4 months. There was high-fidelity use (median WHOBARS score 5.0 of 7; use of basic safety processes ranged from 85.0 to 99.0 per cent), and high penetration shown by a significant improvement in hospital safety culture (adapted Human Factors Attitude Questionnaire scores of 76.7, 81.1 and 82.2 per cent before, and at 4 and 12–18 months after training respectively; $P < 0.001$). Acceptability, adoption, appropriateness and feasibility scored 9.6–9.8 of 10. This approach incorporated 31 of 36 CFIR implementation constructs successfully.

Conclusion: This study shows successfully sustained nationwide checklist implementation using a validated implementation framework.

Presented to the West African College of Surgeons Meeting, Banjul, Gambia, March 2018, and the Group of Anaesthetists in Training Meeting, Glasgow, UK, July 2018; published in abstract form as *Anaesthesia* 2018; 73(Suppl S3): 14



Paper accepted 1 October 2018

Published online in Wiley Online Library (www.bjs.co.uk). DOI: 10.1002/bjs.11034

Introduction

Surgical patient safety has been a major preoccupation of the profession for the past two decades^{1,2}. Much progress has been made in identifying what causes errors, and understanding the relationship between safety culture and patient outcome. However, these advances are yet to materialize in low- and middle-income countries (LMICs), where surgical safety remains poor and outcomes

are significantly worse than those in high-income countries (HICs); this creates very significant disparity and inequity worldwide^{3–6}. Although some of the differences between LMICs and HICs may be explained by lack of infrastructure⁷, low workforce density⁸, insufficient surgical volume^{9,10}, catastrophic surgical costs¹¹ and reluctance to seek care^{12–14}, numerous affordable, life-saving interventions do exist¹⁵. However, too often interventions

that are effective in small-scale pilot studies fail to live up to expectations when rolled out in national strategies, or fail to translate successfully from HICs to LMICs. Therefore, one of the greatest challenges facing the global surgical community is how to take proven perioperative safety and quality improvement interventions and implement them successfully at scale in LMICs.

The World Health Organization (WHO) Surgical Safety Checklist^{16,17} is a good candidate for LMIC scale-up. The checklist has been shown repeatedly to improve surgical outcomes, although poor implementation can negate the benefits^{18–22}. Checklist use is widespread in HICs, but less so in LMICs where evidence of successful implementation is generally limited to small or single-centre studies^{23–27}. In 2011, the WHO unsuccessfully attempted nationwide checklist implementation in 15 African countries²⁸, but after 1 year only ten hospitals in 15 countries had started using the checklist. Thus, the checklist is a good example of an ‘implementation gap’, waiting to be addressed in surgical care in LMICs. Implementation research is a form of health policy and systems research that can be used to study and support the scale-up of quality improvement interventions into health systems at the national level²⁹. Recent research in Madagascar^{30,31} has shown that nationwide checklist implementation is feasible and results in high-fidelity checklist use.

This study used implementation science principles to evaluate nationwide checklist implementation in Benin. Based on studies in Madagascar³¹ and England³², the hypothesis was that sustained checklist use would be reported by at least 50 per cent of participants at 12–18 months after introduction. The aim was to measure the sustainability of checklist use; and to evaluate the acceptability, adoption, appropriateness, feasibility and fidelity of nationwide checklist implementation, including penetration of the checklist into operating room culture and its impact on individual staff.

Methods

The Benin Ministry of Health, Mercy Ships Institutional Review Board and King’s College London Research Ethics Committee approved the study. All participants gave voluntary written informed consent to participate and no incentive payments were made. The study is reported in accordance with the STROBE guidelines³³.

The study was a collaboration between the non-governmental organization Mercy Ships, the Benin Ministry of Health and the academic institution, King’s College London. Mercy Ships visits countries at the invitation of the Head of State, typically spending 10 months delivering

Table 1 Implementation outcomes: definitions and measurement tools used in present study

Outcome	Definition	Measurement tool used
Acceptability	Perception among stakeholders that the intervention is agreeable	Questionnaire 2
Adoption	Willingness to start employing the intervention	Questionnaire 2
Appropriateness	Perception among stakeholders of the fit and relevance of the intervention to the local context	Questionnaire 2
Cost	Cost impact of the implementation effort	–
Feasibility	Extent to which an intervention can be carried out successfully	Questionnaire 2 Focus group
Fidelity	Degree to which the intervention is implemented as originally intended	Questionnaire 1 WHOBARS
Penetration	Integration of an intervention within a service system	Questionnaire 1 HFAQ Focus group
Sustainability	Extent to which a newly implemented intervention is maintained within a service setting	Questionnaire 1 Focus group

WHOBARS, WHO Behaviourally Anchored Rating Scale; HFAQ, Human Factors Attitude Questionnaire.

free operations and training in partnership with the Ministry of Health, aiming to strengthen the surgical ecosystem, and advocate access to safe affordable timely surgery and development of national surgical plans^{34–37}.

Standardized implementation outcomes

Eight distinct implementation outcomes are the standard for evaluating the effectiveness of implementation of a clinical intervention (*Table 1*)³⁸. This framework forms the basis for the primary (sustainability) and secondary outcomes. These implementation outcomes fit within Proctor’s broader framework of distinct intervention outcomes for clinical interventions in health: service system, clinical/patient and implementation³⁹ (*Table S1*, supporting information).

Consolidated Framework for Implementation Research

The Consolidated Framework for Implementation Research (CFIR)⁴⁰ was developed to consolidate theories associated with successful implementation outcomes. The CFIR was recently used to describe the process of surgical system strengthening through the development of national surgical plans⁴¹. The CFIR comprises five domains shown to influence how clinical interventions become embedded into health systems: intervention characteristics; outer setting; inner setting; characteristics of

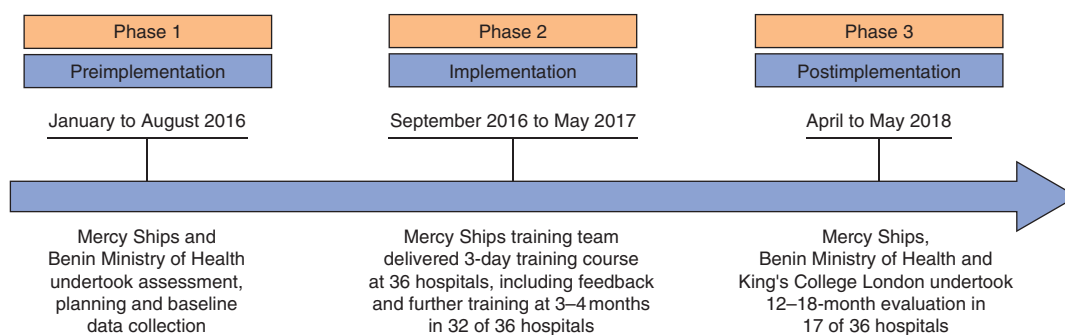


Fig. 1 The three phases of the implementation process

the individuals involved; and the implementation process. It forms the basis for the implementation strategy in Benin (Table S2, supporting information), and helps examine why the approach to nationwide checklist implementation is successful and highlights principles that may be useful for scale-up of other quality improvement initiatives in LMICs.

Study design

A longitudinal embedded mixed-methods design⁴² was used to address the research objectives, whereby quantitative data were the primary focus, backed up by qualitative data, to explore what worked and why. The study consisted of three distinct phases (Fig. 1; Table S3, supporting information).

Phase 1: preimplementation assessment and planning

Mercy Ships, in partnership with the Benin Ministry of Health, purposively selected 36 hospitals that represented the majority of government surgical hospitals. Each hospital director received written introductory information, and completed a baseline questionnaire. This formed the basis of the initial feasibility assessment, allowed the training team to anticipate potential implementation barriers, and encouraged hospital leadership engagement.

Phase 2: implementation process

The initial 3-day checklist training course was undertaken in each hospital. Hospital directors invited the entire surgical team to attend. This included surgical and anaesthesia providers, nurses and any other perioperative staff members (such as nursing aides and surgical assistants). Hospitals were requested not to schedule elective surgery during the 3-day training. The Ministry of Health and hospital directors considered that, as elective surgery did not happen consistently every day, there would be sufficient capacity to reschedule elective surgery around

the training course. The course timetable was adapted to accommodate regular hospital activities such as morning rounds, and any emergency operations performed were often used to facilitate checklist training in real time.

Three to five faculty members taught the course, after receiving training from the lead author. A Beninois generalist doctor and Togolese final year medical student (studying in Benin) were always present together with at least one UK surgical or anaesthetic trainee, and an operating room nurse. The course highlights the evidence base for the checklist; uses multidisciplinary simulation and discussion to adapt the checklist to the host hospital environment; teaches specific skills such as counting needles, swabs and instruments; and donates pulse oximeters if required. At the end of the course, the hospital-specific checklist, including surgical counting tool, is given to the surgical team and presented to the hospital director. A more detailed course description and timetable has been published previously^{30,31}.

At the end of the training, course participants were invited to join a Benin Checklist WhatsApp group. The purpose of the group was to develop a communication network between hospitals, and promote a positive checklist culture and climate for implementation through shared experiences. Formal evaluation of the WhatsApp group was not planned *a priori*, but hospital staff used the WhatsApp group mainly to share experiences as well as to ask questions of the wider group; it therefore functioned as an easily accessible peer support group. Questions were responded to by both other hospital staff and the training team, and inappropriate use of the group (such as for selling motorbikes) was discouraged. Informal telephone follow-up occurred at 6 weeks to one or two members of hospital staff identified during the training as influential and enthusiastic about checklist implementation. Site visits were conducted 3–4 months after the course to evaluate early participant experiences and organizational change through surveys and focus groups. This allowed

feedback and an opportunity to address ongoing implementation challenges at either individual or hospital leadership level.

Phase 3: longitudinal evaluation at 12–18 months

Over a 4-week interval, 17 of the original 36 hospitals were revisited for a day at a time. To achieve a representative sample, hospitals were selected purposively based on the following criteria: hospital size (large and small); previous performance (good and poor) of checklist implementation at 4 months; and location to maximize the number of hospitals that could be visited within the allocated time interval and budget for evaluation. Hospital directors were contacted by the Ministry of Health to inform them of the visit. The evaluation team consisted of two to five people: three generalist doctors (2 from Benin and 1 from Togo), a British consultant anaesthetist and a British nurse. All team members were trained in using the assessment tools and had experience in checklist evaluation.

Outcome measures

The primary outcome was sustained checklist use at 12–18 months measured by self-reported validated questionnaire^{31,43}.

The secondary outcomes were based on the current standard implementation outcomes framework (*Table 1*). Four specific measurement tools were used as well as focus groups. Questionnaire 1 uses a validated five-point Likert scale response format to measure checklist use^{31,43} and adherence to six basic safety processes¹⁷. Questionnaire 2 uses a visual analogue scale ranging from 1 to 10 to measure the participant's immediate reaction to the training course, based on the Kirkpatrick model of evaluating training courses⁴⁴. The adapted Human Factors Attitude Questionnaire has a five-point Likert scale response format and is used widely to measure organizational safety culture⁴⁵. The WHO Behaviourally Anchored Rating Scale (WHOBARS) is a validated tool that assesses non-technical skills during checklist administration⁴⁶. For each part of the checklist (sign in, time out, sign out), WHOBARS evaluates five domains on a scale from 1 to 7: setting the stage, team engagement, communication activation, communication of problem anticipation, and communication of process completion. Scores are combined and then averaged to give an overall WHOBARS score (range 1–7), with higher scores indicating superior non-technical skills, conducive to high-fidelity checklist application. Because the checklist is thought to exert benefit by improving adherence to safety processes as well as non-technical skills, both safety

processes measured by questionnaire and non-technical skills measured by WHOBARS are needed to evaluate fidelity.

Focus groups used a previously reported discussion guide³⁰ to provide qualitative data on checklist use as experienced by operating room staff.

Data collection

Questionnaire 1 and the adapted Human Factors Attitude Questionnaire were administered before, and at 4 and 12–18 months after training during 1-day hospital visits. Questionnaire 2 was administered immediately after the training course to determine the participants' reaction to the course (acceptability, adoption, appropriateness) and feasibility of implementation in their hospital.

WHOBARS assessments were made during the 12–18-month hospital visit using direct observations of checklist administration in real time in the operating room. If no surgery occurred during the evaluation visit, simulation was used instead to measure WHOBARS. During simulation, participants were asked to adopt their usual professional role. One focus group was held per hospital at 3–4 months and again at 12–18 months with all available staff. All focus groups took place in the participants' hospital and lasted 20–60 min. Focus groups at 3–4 months were conducted in English with French concurrent translation, using a facilitator who moderated the discussion, and primary and secondary scribes who took notes of the discussion. These notes were used to provide feedback, further training and problem-solving as required to improve checklist implementation. At 12–18 months, focus groups were conducted in French, tape recorded and transcribed verbatim.

Data analysis

Human Factors Attitude Questionnaire Likert responses were scored on a scale from 1 to 5; values of negatively worded questions were inverted to give equal and positive results so that the overall score could be calculated and given as a percentage. Open questionnaire responses and all focus group data were grouped by category or question in Excel[®] (Microsoft, Redmond, Washington, USA) and then analysed manually by the investigators using inductive thematic analysis^{47,48}. Inductive thematic analysis was used to analyse key changes in practice and culture, and facilitators and barriers associated with implementation. Important topics were identified and grouped into related themes. No software was used for the qualitative analyses.

Table 2 Hospital and participant demographics

	Initial training	3–4-month evaluation	12–18-month evaluation
Hospital			
No. of hospitals	36	32	17
Size (no. of hospital beds)			
Median	100	100	120
Range	52–898	52–898	70–898
i.q.r.	82–131	70–120	110–224
Surgical volume (no. of operations per month)			
Median	80	76	148
Range	20–288	20–288	35–288
i.q.r.	47–148	50–150	75–183
Participants			
No. of participants	638	192	110
No. from whom questionnaire data collected	543	189	104
Surgeons	79 (14.5)	37 (19.6)	26 (25.0)
Anaesthetists	89 (16.4)	39 (20.6)	22 (21.2)
Surgical assistants	48 (8.8)	23 (12.2)	14 (13.5)
Nurses	193 (35.5)	44 (23.3)	23 (22.1)
Other health aides	127 (23.4)	40 (21.2)	15 (14.4)
Not recorded	7 (1.3)	6 (3.2)	4 (3.8)
No. of participants per hospital			
Median	11	7	5
Range	2–62	2–23	1–12
i.q.r.	3–19	6–12	5–7

Values in parentheses are percentages.

Table 3 Frequency of checklist use before, and 4 and 12–18 months after checklist training

	No. of respondents	Always, in full	Always, in part	Sometimes	Occasionally	Never
Before training	543	100 (18.4)	69 (12.7)	85 (15.7)	59 (10.9)	230 (42.4)
4 months	178	97 (54.5)	61 (34.3)	13 (7.3)	4 (2.2)	3 (1.7)
12–18 months	100	55 (55.0)	31 (31.0)	9 (9.0)	4 (4.0)	1 (1.0)

Values in parentheses are percentages.

Statistical analysis

Simple descriptive statistics were used to explore the primary (sustainability) and secondary (state) outcomes. Survey scores at each time interval were compared using a *t* test and $P < 0.050$ was considered significant. All calculations were performed in Excel®.

Results

The median interval from initial checklist training to longitudinal sustainability assessment was 17 (range 14–18, i.q.r. 15–18) months.

Details of the hospital and participant demographics for the original training, 3–4-month and 12–18-month evaluations are shown in *Table 2*. Four hospitals were omitted at the 3–4-month follow-up owing to security concerns and consequent change to the Mercy Ships zone of safe travel.

Sustainability

Before training, 31.1 per cent of participants (169 of 543) reported using at least part of the WHO checklist all the time. This increased to 88.8 per cent (158 of 178) at 4 months, and at 12–18 months, 86.0 per cent of participants (86 of 100) reported sustained checklist use (*Table 3*). In addition, the WhatsApp group continued at 18 months with commentary only from Benin operating room staff.

Acceptability, adoption and appropriateness

Immediately after the course, 543 of 638 participants completed Questionnaire 2. They rated the acceptability of the training course, willingness to adopt the checklist and appropriateness of the checklist to fit the host environment as 9.8 (95 per cent c.i. 9.7 to 9.8), 9.7 (9.6 to 9.8) and 9.6 (9.5 to 9.7) of 10 respectively. Open questionnaire responses described the training as ‘very dynamic

Table 4 Frequency of use of the six basic safety processes before, and 4 and 12–18 months after checklist training

	No. of respondents	Always, in full	Most of the time	Sometimes	Occasionally	Never
Process 1						
Before training	543	427 (78.6)	68 (12.5)	24 (4.4)	16 (2.9)	8 (1.5)
4 months	178	147 (82.6)	23 (12.9)	5 (2.8)	2 (1.1)	1 (0.6)
12–18 months	100	77 (77.0)	19 (19.0)	2 (2.0)	2 (2.0)	0 (0)
Process 2						
Before training	543	284 (52.3)	142 (26.2)	68 (12.5)	38 (7.0)	11 (2.0)
4 months	178	103 (57.9)	57 (32.0)	12 (6.7)	6 (3.4)	0 (0)
12–18 months	100	68 (68.0)	23 (23.0)	6 (6.0)	3 (3.0)	0 (0)
Process 3						
Before training	543	293 (54.0)	155 (28.5)	59 (10.9)	24 (4.4)	12 (2.2)
4 months	178	92 (51.7)	55 (30.9)	21 (11.8)	6 (3.4)	4 (2.2)
12–18 months	100	58 (58.0)	27 (27.0)	10 (10.0)	1 (1.0)	4 (4.0)
Process 4						
Before training	543	470 (86.6)	54 (9.9)	10 (1.8)	4 (0.7)	5 (0.9)
4 months	178	149 (83.7)	26 (14.6)	0 (0)	1 (0.6)	2 (1.1)
12–18 months	100	97 (97.0)	2 (2.0)	1 (1.0)	0 (0)	0 (0)
Process 5						
Before training	543	156 (28.7)	124 (22.8)	119 (21.9)	86 (15.8)	58 (10.7)
4 months	178	100 (56.2)	57 (32.0)	16 (9.0)	5 (2.8)	0 (0)
12–18 months	100	57 (57.0)	30 (30.0)	9 (9.0)	4 (4.0)	0 (0)
Process 6						
Before training	543	236 (43.5)	98 (18.0)	83 (15.3)	55 (10.1)	71 (13.1)
4 months	178	127 (71.3)	39 (21.9)	9 (5.1)	3 (1.7)	0 (0)
12–18 months	100	67 (67.0)	23 (23.0)	5 (5.0)	5 (5.0)	0 (0)

Values in parentheses are percentages. Basic safety processes: 1, identity of the patient, and type and site of surgery verified; 2, evaluation of risk of difficult intubation; 3, evaluation of risk of large blood loss; 4, use of a pulse oximeter; 5, prophylactic antibiotics given before surgical skin incision; 6, needles/sponges/instruments counted before and after surgery.

and engaging' and that 'everyone got to participate and be a part of it'. Important learning points reported by participants were 'correct counting of compresses, needles and instruments' and 'avoiding accidents such as forgetting materials in the patient's abdomen'. A willingness to adopt the checklist was shown by comments such as: 'we will check our equipment before starting each surgery', 'we will do a correct surgical count before and after every surgery' and 'we will help to put the patient at ease and communicate better with everyone'.

Feasibility

Immediately after the course, participants who completed Questionnaire 2 rated the feasibility of checklist implementation to improve surgical safety as 9.7 (95 per cent c.i. 9.7 to 9.8) of 10. The commonest changes participants anticipated making in order to implement the checklist were checking equipment before starting surgery and performing a surgical count; the commonest behaviour changes anticipated were improved teamwork, communication and ensuring that patients would feel more at ease. Anticipated barriers to implementation were resistance of senior colleagues who had not attended the training, lack of staff and the perception that sometimes the checklist can

take a long time so it might be difficult to administer during emergency procedures. Pulse oximeters were reported as present and used routinely by 96.5 per cent of participants (524 of 543) before checklist implementation. During the implementation process, seven pulse oximeters were donated to a total of five hospitals, to ensure that each recovery area in every hospital had a sufficient number available. At 4 and 18 months after implementation, pulse oximeters were used routinely by 98.3 per cent (175 of 178) and 99.0 per cent (99 of 100) respectively. The feasibility of nationwide checklist implementation was backed up by data on the primary outcome measure; participants reported an increase in checklist use from 31.1 per cent before training to 88.8 per cent at 4 months, sustained at 86.0 per cent at 17 months.

Fidelity

After 12–18 months, of the six basic safety processes associated with checklist use, pulse oximetry, and verifying the patient's identity and type of surgery were the most frequently done (99.0 per cent (99 of 100) and 96.0 per cent (96 of 100) respectively), whereas assessment of the risk of blood loss was the least frequently carried out (85.0 per cent, 85 of 100) (Table 4). From the focus groups it

Table 5 Mean hospital WHOBARS scores and method of evaluation

Hospital	Method of evaluation	WHOBARS score			Overall
		Sign in	Time out	Sign out	
D	Real-time procedure in OR	6.1	5.8	1.4	4.3
F	Real-time procedure in OR	4.2	4.2	3.8	4.1
I	Real-time procedure in OR	–	–	5.6	5.6
M	Real-time procedure in OR	4.8	4.8	4.2	4.6
O	Real-time procedure in OR	6.0	6.0	4.0	5.3
P	Real-time procedure in OR	1.0	1.0	1.0	1.0
A	Simulation	6.0	6.0	6.0	6.0
C	Simulation	6.0	5.9	6.0	6.0
H	Simulation	4.4	4.9	3.2	4.2
J	Simulation	6.8	6.8	6.6	6.7
K	Simulation	3.6	4.4	3	3.7
Q	Simulation	6.0	6.0	6.0	6.0
B	–	–	–	–	–
E	–	–	–	–	–
G	–	–	–	–	–
L	–	–	–	–	–
N	–	–	–	–	–

The WHO Behaviourally Anchored Rating Scale (WHOBARS) ranged from 1 to 7. OR, operating room; –, not done.

was clear that some of the improvement in process may have involved policies that extended beyond the operating room. Some hospitals reported new processes introduced by operating room managers such as instrument trays having a standard set of instruments (for example for caesarean section), and replacing compresses and sponges from single large pots into smaller pots with batches of five to make counting easier.

WHOBARS assessment of team behaviour during checklist administration occurred in 12 of 17 hospitals, and was conducted on real procedures in six hospitals and using simulation in six hospitals (Table 5). WHOBARS assessment was not undertaken in a total of five hospitals owing to refusal of consent in two and lack of time in three hospitals. The median overall WHOBARS score was 5.0 (range 1.0–6.7, i.q.r. 4.1– 6.0). In the study data set, WHOBARS score did not correlate with hospital-level checklist use ($r = 0.44$, $P = 0.151$).

Penetration

The improved Human Factors Attitude Questionnaire scores showed penetration of the checklist into hospital safety culture. Before checklist implementation, mean survey scores were 76.7 (95 per cent c.i. 75.9 to 77.6) per cent among 543 participants, which increased significantly to 81.1 (79.6 to 82.6) per cent among 189 participants at 4 months ($P < 0.001$). At 12–18 months, the improved scores were sustained at 82.2 (80.7 to 85.0) per cent among 104 participants ($P = 0.185$). Increasing Human Factors Attitude Questionnaire score correlated weakly

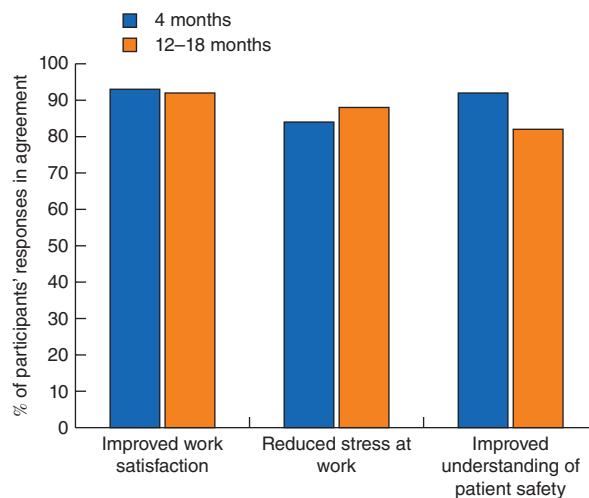


Fig. 2 Impact of checklist use at 12–18 months on individual well-being and understanding of patient safety

and positively with hospital-level checklist use ($r = 0.57$, $P = 0.052$).

Fig. 2 shows the penetration of the checklist into self-reported individual experience of well-being and understanding of patient safety.

Thematic analysis of focus group transcripts showed that the common features for hospitals successfully implementing the checklist were enthusiastic surgical leadership (11 of 17 hospitals), supportive hospital administrative leadership (10 of 17), and seeing the value of the checklist for both patient safety, and improved teamwork and communication

(9 of 17). These hospitals all described checklist use as systematic (10 of 17) and as something that the whole team liked because they found it helpful (10 of 17). In hospitals where checklist implementation was unsuccessful, staff viewed the checklist as irrelevant or a waste of time and leadership appeared arrogant. Further details and speech quotations are shown in *Table S4* (supporting information).

Discussion

This study of nationwide WHO checklist implementation in Benin found that 12–18 months after a 3-day training course, 86.0 per cent of participants reported sustained checklist use compared with 31.1 per cent before implementation. Because outcomes improve as checklist compliance improves, compliance rates of less than 100 per cent can still be clinically significant^{18,19,49}. Indeed, wide variation in compliance rates (39–100 per cent) has been shown in England and New Zealand^{32,50}, and the present results showed a similar variation in individual checklist items and WHOBARs scores. These results are comparable to those of a previous study in Madagascar^{30,31}, which used the same 3-day multidisciplinary training model for nationwide checklist implementation. This indicates that the present model of checklist scale-up is transferable between countries.

Use of the CFIR⁴⁰ to guide the implementation strategy helps explain this successful result. The checklist rates highly in the CFIR constructs known to affect successful implementation (31 of 36 constructs). The checklist is recommended by the WHO, has a strong evidence base, is simple to use and the benefits are clearly seen. A key CFIR construct related to the intervention characteristics is the idea of a ‘central core and an adaptable periphery’. The checklist course design embraced this construct as the implementation team ensured that each hospital kept the core components (the 6 basic safety steps) but adapted the other checklist items to suit the local environment. This flexibility is known to overcome implementation barriers³⁰.

In the CFIR the outer and inner setting (outer and inner context of the checklist) have a dynamic interface. This reported study of checklist intervention in Benin sat within the wider context of Ministerial discussions regarding collection of the Lancet Commission on Global Surgery key indicators and developing a national surgical plan⁵¹, and used both a ‘top-down’ and a ‘bottom-up’ approach. There was a detailed planning process in collaboration with the Ministry of Health and hospital directors, and strong follow-up and feedback with individuals through telephone calls, WhatsApp group and in-person hospital visits. This in-depth stakeholder engagement at multiple levels was

likely an important component to success and the strong relationships facilitated troubleshooting at the 4-month feedback visit. In two hospitals, use of the checklist was seen to be difficult at 4 months and a focus group discussion resulted in a stakeholder meeting with the hospital director. Changes were recommended and both hospitals reported high-fidelity checklist use at 18 months. Other features of the inner setting, such as structural characteristics (hospital size and surgical volume), have not been associated with implementation success in other countries³¹. It is likely that the CFIR constructs of culture, implementation climate and readiness are more important. The implementation strategy in the present study facilitated these aspects by creating: a multidisciplinary learning environment through workshop style teaching, simulation and real-life practice; psychological ownership by permitting locally adapted hospital-specific checklists; providing goals and feedback through the 4-month evaluation; an ongoing network of communication through the WhatsApp group; and continuous multilevel stakeholder engagement.

The characteristics of individuals (CFIR domain 4) affect the success of intervention scale-up because individuals wield power and influence. The results reported in this paper show that surgical enthusiasm for the checklist is one of the most important factors in successful implementation. However, it requires more than one enthusiastic individual to sustain checklist use. Self-efficacy and motivation is another important aspect of sustained change, demonstrated in this study by the initiative shown in seven of 17 hospitals that decided to photocopy the checklist and counting sheet and place them in the patients’ medical record to promote implementation.

The 2015 Lancet Commission on Global Surgery⁵¹ and a plethora of subsequent publications continue to describe a global lack of access to safe, affordable surgical care⁵². However, the surgical community must move beyond identifying problems to finding real-life solutions⁵². To date, little has been published detailing strategies for scale-up of quality improvements in LMICs. The authors suggest that this is because the issues facing LMICs pose problems of a fundamental nature at the systems level: what safety and quality improvement interventions the system can support, absorb and sustain. Lack of implementation at the systems level may be why well evidenced improvement solutions in HIC surgical systems have previously fared poorly in LMICs²⁸. Using the CFIR approach, the intervention, setting and individuals act as parameters within which the implementer must act but has little control over. Therefore, implementers must work with all the stakeholders to design implementation processes that overcome known barriers at multiple levels. This study involved a

collaboration of three major stakeholders: a Ministry of Health, a non-governmental organization and an academic institution, as well as stakeholders at the hospital and individual operating room worker level. This collaborative process, using a marriage of implementation science and health systems strengthening approaches, is likely to be important in scaling up other surgical care improvements in LMICs.

This study has limitations. The evaluation team was not completely independent of the implementation team, but having a member of the evaluation team who is known to participants has been shown to reduce responder bias in qualitative studies^{53,54}. No assessment of clinical outcomes was made, but where the intervention is well evidenced to improve clinical outcome, it is not necessary to continue asking 'does it work?'; thus the focus was 'how can we make it work at scale in real life in a LMIC?'. The study lacked control sites because of a limited time frame of 10 months within which to conduct a national implementation at 36 hospitals, including 4-month follow-up visits to each site (phase 2). This limitation could be addressed in future national studies by using a step wedge design. The present study involved both the implementation process and a research component, but it was not possible to evaluate whether the same result could have been achieved had the intervention been stripped of the research components (surveys, focus groups, observations with WHO-BARS). The question thus remains to what extent research efforts in LMICs skew findings on implementability and sustainability of surgical improvement interventions. Much of the data were obtained by self-reported questionnaires or focus groups and are therefore open to subjective bias: under-reporting in the hope of obtaining further training and support, or over-reporting to gain favour from non-governmental organization or government stakeholders. The Human Factors Attitude Questionnaire scores were statistically different but this may not indicate a meaningful clinical difference. Both the checklist questionnaire and the Human Factors Attitude Questionnaire may not have sampled the same participants in repeated surveys and the numbers declined progressively. Measurement of WHOBARs scores throughout the study was limited because of lack of consent to participate (2 hospitals), and lack of time and resource for the lead investigator to visit a further three hospitals to collect WHOBARs data from the entire cohort. This reflects pragmatic limitations in data collection with limited funding in a LMIC setting, but may have had an impact on the behavioural data collected through WHOBARs. The WHOBARs data (non-technical skills around checklist implementation) collected pragmatically in a mixture of real operating

room and simulated procedures may have differed between the two settings. On balance, the mixed-methods design and breadth of quantitative and qualitative data aimed to reduce bias by allowing triangulation of the results. The study comprised high-intensity implementation with 3-day in-person training, operating room stand down, multiple site visits, surveys, telephone calls, WhatsApp group and focus groups. This kind of implementation performed at scale in LMICs requires significant human and financial resources, as well as requisite expertise, which were not quantified in this study.

Acknowledgements

The authors acknowledge the following Benin Ministry of Health officials for logistical help during different phases of the study: S. E. Gbedo (Conseiller Technique au Partenariat Sanitaire); M. Amoussou-Guenou (Directeur Direction des Hôpitaux); S. Gbenou (Directeur Adjoint Direction des Hôpitaux); T. A. Ahojo (Directrice Sécurité des Patients); A. A. Kokou (Cellule de Sécurité des Patients et de la Gestion des Risques); L. Affovehoude (Chef Division Sécurité et Satisfaction des Patients); and E. Tokplonou (Chef Service Assurance Qualité et prestations de soins à la direction Nationale des Hôpitaux). They also thank the following Mercy Ships volunteers for help with checklist course teaching and data collection: M. Nicholson, K. Harvey-Kelly, A. P. Herbert, G. Nickols and J. Tay.

Because of the sensitive nature of the data collected for this study, requests to access the data set from qualified researchers trained in human subject confidentiality protocols may be sent to the corresponding author and are subject to approval by Mercy Ships Institutional Review Board and the Benin Ministry of Health.

S.R. and N.S. are supported by the National Institute for Health Research (NIHR) Collaboration for Leadership in Applied Health Research and Care (CLAHRC) South London at King's College Hospital NHS Foundation Trust. S.R. is supported by the NIHR through a Knowledge Mobilization Fellowship. N.S. is a member of King's Improvement Science, which is part of the NIHR CLAHRC South London and comprises a specialist team of improvement scientists and senior researchers based at King's College London. Its work is funded by King's Health Partners (Guy's and St Thomas' NHS Foundation Trust, King's College Hospital NHS Foundation Trust, King's College London, and South London and Maudsley NHS Foundation Trust), Guy's and St Thomas' Charity, the Maudsley Charity and the Health Foundation. A.J.M.L. and N.S. are supported by the NIHR Global

Health Research Unit on Health System Strengthening in Sub-Saharan Africa, King's College London (GHRU 16/136/54). A.J.M.L. and N.S. are further supported by the ASPIRES (Antibiotic use across Surgical Pathways – Investigating, Redesigning and Evaluating Systems) research programme in LMICs, funded by the Economic and Social Research Council.

The views expressed are those of the authors and not necessarily those of the National Health Service, the NIHR or the Department of Health and Social Care.

N.S. is the director of London Safety and Training Solutions, which offers training in patient safety, implementation solutions and human factors to healthcare organizations.

Disclosure: The authors declare no other conflict of interest.

References

- Institute of Medicine. *To Err Is Human: Building a Safer Health System*. Kohn LT, Corrigan JM, Donaldson MS (eds). National Academies Press: Washington, 2000.
- Weiser TG, Haynes AB. Ten years of the Surgical Safety Checklist. *Br J Surg* 2018; **105**: 927–929.
- GlobalSurg Collaborative. Determinants of morbidity and mortality following emergency abdominal surgery in children in low-income and middle-income countries. *BMJ Glob Health* 2016; **1**: e000091.
- GlobalSurg Collaborative. Mortality of emergency abdominal surgery in high-, middle- and low-income countries. *Br J Surg* 2016; **103**: 971–988.
- Biccard BM, Madiba TE, Kluyts HL, Munlemvo DM, Madzimbamuto FD, Basenero A *et al.*; African Surgical Outcomes Study (ASOS) investigators. Perioperative patient outcomes in the African Surgical Outcomes Study: a 7-day prospective observational cohort study. *Lancet* 2018; **391**: 1589–1598.
- GlobalSurg Collaborative. Surgical site infection after gastrointestinal surgery in high-income, middle-income, and low-income countries: a prospective, international, multicentre cohort study. *Lancet Infect Dis* 2018; **18**: 516–525.
- Hsia RY, Mbembati NA, Macfarlane S, Kruk ME. Access to emergency and surgical care in sub-Saharan Africa: the infrastructure gap. *Health Policy Plan* 2012; **27**: 234–244.
- Holmer H, Shrimme MG, Riesel JN, Meara JG, Hagander L. Towards closing the gap of the global surgeon, anaesthesiologist, and obstetrician workforce: thresholds and projections towards 2030. *Lancet* 2015; **385**(Suppl 2): S40.
- Rose J, Weiser TG, Hider P, Wilson L, Gruen RL, Bickler SW. Estimated need for surgery worldwide based on prevalence of diseases: a modelling strategy for the WHO Global Health Estimate. *Lancet Glob Health* 2015; **3**(Suppl 2): S13–S20.
- Weiser TG, Haynes AB, Molina G, Lipsitz SR, Esquivel MM, Uribe-Leitz T *et al.* Estimate of the global volume of surgery in 2012: an assessment supporting improved health outcomes. *Lancet* 2015; **385**(Suppl 2): S11.
- Shrimme MG, Dare AJ, Alkire BC, O'Neill K, Meara JG. Catastrophic expenditure to pay for surgery worldwide: a modelling study. *Lancet Glob Health* 2015; **3**(Suppl 2): S38–S44.
- Nagarajan N, Gupta S, Shrestha S, Varadaraj V, Devkota S, Ranjit A *et al.* Unmet surgical needs in children: a household survey in Nepal. *Pediatr Surg Int* 2015; **31**: 389–395.
- Zhang XJ, Jhanji V, Leung CK, Li EY, Liu Y, Zheng C *et al.* Barriers for poor cataract surgery uptake among patients with operable cataract in a program of outreach screening and low-cost surgery in rural China. *Ophthalmic Epidemiol* 2014; **21**: 153–160.
- Nguyen K, Bhattacharya SD, Maloney MJ, Figueroa L, Taicher BM, Ross S *et al.* Self-reported barriers to pediatric surgical care in Guatemala. *Am Surg* 2013; **79**: 885–888.
- Peters DH, El-Saharty S, Siadat B, Janovsky K, Vujicic M. *Improving Health Service Delivery in Developing Countries: From Evidence to Action*. World Bank: Washington, 2009.
- WHO. *WHO Guidelines for Safe Surgery*. WHO Press: Geneva, 2009.
- Haynes AB, Weiser TG, Berry WR, Lipsitz SR, Breizat AS, Dellinger P *et al.* A surgical safety checklist to reduce morbidity and mortality in a global population. *N Engl J Med* 2009; **360**: 491–499.
- Bergs J, Hellings J, Cleemput I, Zurel Ö, De Troyer V, Van Hiel M *et al.* Systematic review and meta-analysis of the effect of the World Health Organization surgical safety checklist on postoperative complications. *Br J Surg* 2014; **101**: 150–158.
- Haugen AS, Söfteland E, Almeland SK, Sevdalis N, Vonen B, Eide GE *et al.* Effect of the World Health Organization checklist on patient outcomes: a stepped wedge cluster randomized controlled trial. *Ann Surg* 2015; **261**: 821–828.
- Haugen AS, Wæhle HV, Almeland SK, Harthug S, Sevdalis N, Eide GE *et al.* Causal analysis of World Health Organization's Surgical Safety Checklist implementation quality and impact on care processes and patient outcomes: secondary analysis from a large stepped wedge cluster randomized controlled trial in Norway. *Ann Surg* 2017; [Epub ahead of print].
- Haynes AB, Edmondson L, Lipsitz SR, Molina G, Neville BA, Singer SJ *et al.* Mortality trends after a voluntary checklist-based surgical safety collaborative. *Ann Surg* 2017; **266**: 923–929.
- Levy SM, Senter CE, Hawkins RB, Zhao JY, Doody K, Kao LS *et al.* Implementing a surgical checklist: more than checking a box. *Surgery* 2012; **152**: 331–336.
- Kwok AC, Funk LM, Baltaga R, Lipsitz SR, Merry AF, Dziekan G *et al.* Implementation of the World Health Organization surgical safety checklist, including introduction of pulse oximetry, in a resource-limited setting. *Ann Surg* 2013; **257**: 633–639.

- 24 Kasatpibal N, Senaratana W, Chitreecheur J, Chotirosniramit N, Pakvipas P, Junthasopeepun P. Implementation of the World Health Organization surgical safety checklist at a university hospital in Thailand. *Surg Infect (Larchmt)* 2012; **13**: 50–56.
- 25 Lilaonitkul M, Kwikiriza A, Ttendo S, Kiwanuka J, Munyarungero E, Walker IA *et al.* Implementation of the WHO Surgical Safety Checklist and surgical swab and instrument counts at a regional referral hospital in Uganda – a quality improvement project. *Anaesthesia* 2015; **70**: 1345–1355.
- 26 Bashford T, Reshamwalla S, McAuley J, Allen NH, McNatt Z, Gebremedhen YD. Implementation of the WHO Surgical Safety Checklist in an Ethiopian referral hospital. *Patient Saf Surg* 2014; **8**: 16.
- 27 Yuan CT, Walsh D, Tomarken JL, Alpern R, Shakpeh J, Bradley EH. Incorporating the World Health Organization Surgical Safety Checklist into practice at two hospitals in Liberia. *Jt Comm J Qual Patient Saf* 2012; **38**: 254–260.
- 28 Kariyoi PC, Hightower J, Ndiokubwayo JB, Tumusiime P, Mwikisa C. Challenges facing the introduction of the WHO surgical safety checklist: a short experience in African countries. *Afr Health Monit* 2013; **16**: 36–39.
- 29 Peters DT, Nhan T, Adam T; Alliance for Health Policy and Systems Research, WHO. *Implementation Research in Health: a Practical Guide*. WHO: Geneva, 2013.
- 30 Close KL, Baxter LS, Ravelojaona VA, Rakotoarison HN, Bruno E, Herbert A *et al.* Overcoming challenges in implementing the WHO Surgical Safety Checklist: lessons learnt from using a checklist training course to facilitate rapid scale up in Madagascar. *BMJ Glob Health* 2017; **2**(Suppl 4): e000430.
- 31 White MC, Baxter LS, Close KL, Ravelojaona VA, Rakotoarison HN, Bruno E *et al.* Evaluation of a countrywide implementation of the world health organisation surgical safety checklist in Madagascar. *PLoS One* 2018; **13**: e0191849.
- 32 Russ S, Rout S, Caris J, Mansell J, Davies R, Mayer E *et al.* Measuring variation in use of the WHO surgical safety checklist in the operating room: a multicenter prospective cross-sectional study. *J Am Coll Surg* 2015; **220**: 1–11.e14.
- 33 EQUATOR Network. *STROBE Reporting Guidelines*; 2007. <http://www.equator-network.org/reporting-guidelines/strobe/> [accessed 19 October 2018].
- 34 Baxter LS, Ravelojaona VA, Rakotoarison HN, Herbert A, Bruno E, Close KL *et al.* An observational assessment of anesthesia capacity in Madagascar as a prerequisite to the development of a National Surgical Plan. *Anesth Analg* 2017; **124**: 2001–2007.
- 35 Bruno E, White MC, Baxter LS, Ravelojaona VA, Rakotoarison HN, Andriamanjato HH *et al.* An evaluation of preparedness, delivery and impact of surgical and anesthesia care in Madagascar: a framework for a National Surgical Plan. *World J Surg* 2017; **41**: 1218–1224.
- 36 Shrimel MG, Hamer M, Mukhopadhyay S, Kunz LM, Claus NH, Randall K *et al.* Effect of removing the barrier of transportation costs on surgical utilisation in Guinea, Madagascar and the Republic of Congo. *BMJ Glob Health* 2017; **2**(Suppl 4): e000434.
- 37 Debas HT. Access to safe surgery. *BMJ Glob Health* 2018; **2**(Suppl 4): e000908.
- 38 Proctor E, Silmere H, Raghavan R, Hovmand P, Aarons G, Bunker A *et al.* Outcomes for implementation research: conceptual distinctions, measurement challenges, and research agenda. *Adm Policy Ment Health* 2011; **38**: 65–76.
- 39 Proctor EK, Landsverk J, Aarons G, Chambers D, Glisson C, Mittman B. Implementation research in mental health services: an emerging science with conceptual, methodological, and training challenges. *Adm Policy Ment Health* 2009; **36**: 24–34.
- 40 Damschroder LJ, Aron DC, Keith RE, Kirsh SR, Alexander JA, Lowery JC. Fostering implementation of health services research findings into practice: a consolidated framework for advancing implementation science. *Implement Sci* 2009; **4**: 50.
- 41 Saluja S, Silverstein A, Mukhopadhyay S, Lin Y, Raykar N, Keshavjee S *et al.* Using the Consolidated Framework for Implementation Research to implement and evaluate national surgical planning. *BMJ Glob Health* 2017; **2**: e000269.
- 42 Creswell JW, Plano Clark VL. *Designing and Conducting Mixed Methods Research*. SAGE Publications: Thousand Oaks, 2011.
- 43 White MC, Peterschmidt J, Callahan J, Fitzgerald JE, Close KL. Interval follow up of a 4-day pilot program to implement the WHO surgical safety checklist at a Congolese hospital. *Global Health* 2017; **13**: 42.
- 44 Kirkpatrick D, Kirkpatrick J. *Evaluating Training Programs*. Berrett-Koehler: San Francisco, 2009.
- 45 University Hospital Southampton. *Understand your Organisation: Adapted Human Factors Attitude Questionnaire, Royal College of Obstetricians and Gynaecologists; Human Factors Attitude Questionnaire (Adapted from Helmreich 1990)*. <https://www.rcog.org.uk/en/careers-training/workplace-workforce-issues/improving-workplace-behaviours-dealing-with-undermining/undermining-toolkit/departmental-and-team-interventions/understand-your-organisation/> [accessed 19 October 2018].
- 46 Devcich DA, Weller J, Mitchell SJ, McLaughlin S, Barker L, Rudolph JW *et al.* A behaviourally anchored rating scale for evaluating the use of the WHO surgical safety checklist: development and initial evaluation of the WHOBARS. *BMJ Qual Saf* 2016; **25**: 778–786.
- 47 Patton MQ. *Qualitative Research & Evaluation Methods: Integrating Theory and Practice*. SAGE Publications: Thousand Oaks, 2015.
- 48 Miles MB, Huberman AM, Saldaña J. *Qualitative Data Analysis: a Methods Sourcebook*. SAGE Publications: Thousand Oaks, 2014.
- 49 van Klei WA, Hoff RG, van Aarnhem EE, Simmermacher RK, Regli LP, Kappen TH *et al.* Effects of the introduction

- of the WHO 'Surgical Safety Checklist' on in-hospital mortality: a cohort study. *Ann Surg* 2012; **255**: 44–49.
- 50 Vogts N, Hannam JA, Merry AF, Mitchell SJ. Compliance and quality in administration of a Surgical Safety Checklist in a tertiary New Zealand hospital. *N Z Med J* 2011; **124**: 48–58.
- 51 Meara JG, Leather AJ, Hagander L, Alkire BC, Alonso N, Ameh EA *et al.* Global Surgery 2030: evidence and solutions for achieving health, welfare, and economic development. *Lancet* 2015; **386**: 569–624.
- 52 Charles AG, Mock C. Advancing global surgery: moving beyond identifying problems to finding solutions. *World J Surg* 2017; **41**: 2979–2980.
- 53 Yusa A, Hynie M, Mitchell S. Utilization of internal evaluation results by community mental health organizations: credibility in different forms. *Eval Program Plann* 2016; **54**: 11–18.
- 54 Conley-Tyler M. A fundamental choice: internal or external evaluation. *Eval J Australasia* 2005; **4**: 3–11.

Supporting information

Additional supporting information can be found online in the Supporting Information section at the end of the article.

Globally, **FIVE BILLION PEOPLE** HAVE NO ACCESS TO SAFE AND AFFORDABLE SURGERY WHEN THEY NEED IT

Mercy Ships operates the largest charity-run hospital ship in the world, taking free medical care and surgery directly to those who need it most.

Help us eradicate the diseases of poverty.

Be part of the crew.



Mercy Ships, The Lighthouse, 12 Meadway Court, Stevenage SG1 2EF
 t +44 (0)1438 727800 | e info@mercyships.org.uk
 www.mercyships.org.uk

volunteering@mercyships.org.uk

