Original article

The Lifebox Surgical Headlight Project: engineering, testing, and field assessment in a resource-constrained setting

N. Starr[1,7], N. Panda[3,4], E. W. Johansen[5], J. A. Forrester[2,7], E. Wayessa[9], D. Rebollo[6], A. August[2], K. Fernandez[7], S. Bitew[7], T. Negussie Mammo[7,10] and T. G. Weiser[2,7,8]

Departments of Surgery, [1]University of California, San Francisco, San Francisco, and [2]Stanford University, Stanford, California, [3]Ariadne Labs, Brigham and Women's Hospital, Harvard T. H. School of Public Health, and [4]Department of Surgery, Massachusetts General Hospital, Boston, and [5]Spark Health Design, Hanover, Massachusetts, and [6]School of Medicine, New York University, New York, USA, [7]Lifebox Foundation, London, and [8]Department of Clinical Surgery, University of Edinburgh, Edinburgh, UK, and Departments of Surgery, [9]Wollega University, Nekempte, and [10]Addis Ababa University, Addis Ababa, Ethiopia

Correspondence to: Dr N. Starr, Department of Surgery, University of California, San Francisco, 505 Parnassus Avenue, S-321, San Francisco, California 94143, USA (e-mail: nichole.starr@ucsf.edu)

Background: **Poor surgical lighting represents a major patient safety issue in low-income countries. This study evaluated device performance and undertook field assessment of high-quality headlights in Ethiopia to identify critical attributes that might improve safety and encourage local use.**

Methods: **Following an open call for submissions (December 2018 to January 2019), medical and technical (non-medical) headlights were identified for controlled specification testing on 14 pre-specified parameters related to light quality/intensity, mounting and battery performance, including standardized illuminance measurements over time. The five highest-performing devices (differential illumination, colour rendering, spot size, mounting and battery duration) were distributed to eight Ethiopian surgeons working in resource-constrained facilities. Surgeons evaluated the devices in operating rooms, and in a comparative session rated each headlight in terms of performance and willingness to purchase.**

Results: **Of 25 submissions, eight headlights (6 surgical and 2 technical) met the criteria for full specification testing. Scores ranged from 8 to 12 (of 14), with differential performance in lighting, mounting and battery domains. Only two headlights met the illuminance parameters of more than 35 000 lux during initial testing, and no headlight satisfied all minimum specifications. Of the five headlights evaluated in Ethiopia, daily operation logbooks noted variability in surgeons' opinions of lighting quality (6–92 per cent) and spot size (0–92 per cent). Qualitative interviews also yielded important feedback, including preference for easy transport. Surgeons sought high quality with price sensitivity (using out-of-pocket funds) and identified the least expensive but high-functioning device as their first choice.**

Conclusion: **No device satisfied all the predetermined specifications, and large price discrepancies were critical factors leading surgeons' choices. The favoured device is undergoing modification by the manufacturer based on design feedback so an affordable, high-quality surgical headlight crafted specifically for the needs of resource-constrained settings can be used to improve surgical safety.**

Presented to the Annual Meeting of the College of Surgeons of East, Central and Southern Africa, Kampala, Uganda, December 2019

Paper accepted 12 May 2020

Published online in Wiley Online Library (www.bjs.co.uk). **DOI:** 10.1002/bjs.11756

Introduction

Adequate illumination of the surgical field is fundamental to the safe conduct of an operation. However, poor surgical lighting in low- and middle-income countries (LMICs) is a pervasive problem. An estimated 125 million operations are undertaken per year in LMICs[1], but up to 30 per cent of hospitals performing these operations will experience intermittent availability of electricity[2]. In a study[3] of surgeons from 39 LMICs, 32 per cent reported delaying or

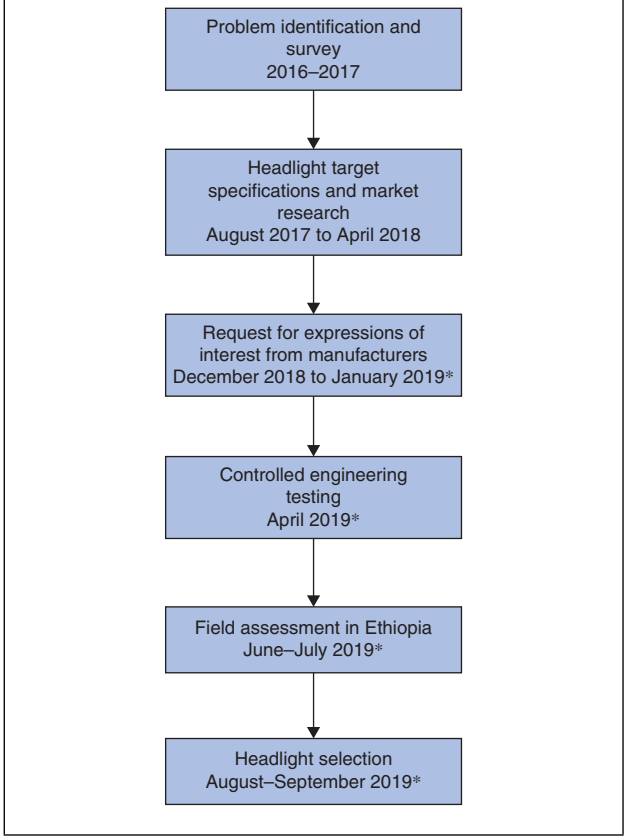
cancelling operations owing to poor lighting, and 47 per cent reported power outages in the operating room at least once per week. A systematic review[4] of 11 sub-Saharan African countries revealed that only 34 per cent of hospitals have consistent access to electricity.

Portable surgical headlights represent a potential solution to shortcomings associated with ineffective overhead operative lighting. Some descriptions of device adaptations in rural and low-resourced surgical environments recommend using camping headlamps or handheld flashlights in lieu of more upscale surgical lighting[5]. However, lighting from recreational headlights, flashlights or mobile phones is neither adequate for visualization of the surgical field nor fit for purpose. Furthermore, recreational headlights often have issues with battery run time and field focus, and can be uncomfortable during longer procedures.

Surgeons in LMICs face challenges in identifying an optimal surgical headlight, as devices are costly (typical retail prices range from US $1200–5000; €1100–4585, exchange rate 24 May 2020) and few opportunities exist for hands-on testing. Certain headlight parameters may be of differential value depending on the procedures performed, resources and infrastructure. Currently available headlight manufacturer specifications are not standardized, as there has been little research comparing surgical headlights. A surgical headlight that is of high quality, rechargeable, and designed with consideration of the specific needs of surgeons in LMICs can provide a more reliable light source during operations, as well as alternative lighting when power fails.

Lifebox, a charity devoted to improving surgical safety worldwide, has launched a programme to identify and distribute a low-cost, robust, fit-for-purpose surgical headlight for use in resource-constrained settings. With input from Lifebox's network of surgeons worldwide, minimum specifications for an appropriately adapted surgical headlight, including light quality, mounting, battery run time, charging and durability, have been established[6]. The specifications also account for exposure to poor air quality, including moisture and dust, and limited access to biomedical engineering services. In the present study, controlled assessments of headlight performance were carried out against these predetermined minimum specifications to identify high-performing devices suitable for surgical use. From this initial screening, devices for field assessment and evaluation in Ethiopia were identified. The aim was to identify surgeons' preferences, the best value device, and specifications that might require further refinement as a means of identifying a fit-for-purpose headlight that will close the lighting gap in surgery and improve the safety of surgical care.

Fig. 1 Lifebox Surgical Headlight Project timeline and project plan



*Processes reported in this article.

Methods

Lifebox Surgical Headlight Project

In 2016, Lifebox began a process to identify an optimal headlight for use in LMICs (*Fig. 1*). As partnerships are an essential element of Lifebox programmes, collaborating with appropriately selected end-users in the medical device design process is of critical importance[7], as human factors assessments while developing or adapting new devices can provide uniquely relevant feedback[8,9]. Human-centred design involves adopting the mindset of the device user, and working through a three-stage process of inspiration, ideation and implementation[10]; the authors' group has been working through this pathway to identify a fit-for-purpose headlight to improve the safety of surgery in resource-poor settings. Key sequential milestones in this process include: assessment of need and potential impact[3]; determination of minimum headlight specifications[6], market research and solicitation of candidate headlights through an open expressions of interest and

a directed request from 100 previously identified manufacturers from 14 different countries[11]; controlled testing of headlights against previously outlined minimum specifications; field assessment of high-performance headlights; and headlight selection and modification plans. These last three tasks are described here.

Identifying headlights for assessment

Minimum specifications for surgical headlight selection were determined previously[6]. After extensive global market research, a request for expressions of interest from headlight manufacturers was issued (*Appendix S1*, supporting information). Both medical and non-medical (hereafter termed 'technical' as many were much more robust than recreational or camping headlights) headlights were included; headlights reliant on mains power were specifically excluded given the ongoing electricity outages and lack of back-up generators in many LMICs. Medical headlights were defined as headlights that were designed biomedically and marketed for use in a clinical setting, whereas technical headlights included headlights used in non-medical activities (for example, skilled trades work such as plumbing or electrical engineering, climbing, caving). Given the ultimate goal of identifying an affordable headlight for distribution at scale, additional screening criteria included: perceived ability to meet predetermined specifications (such as quoted headlight illuminance); quoted wholesale cost of no more than US $600 (€550); potential to work collaboratively with Lifebox; financial viability of the manufacturer of record; and product track record (such as production capacity and product reputation).

Engineering testing protocol

A headlight engineering testing protocol was developed incorporating surgical and human-centred biomedical engineering and design expertise from members of the research team. The protocol was designed to evaluate the performance of surgical and technical headlights against previously articulated specifications related to lighting, mounting and battery features determined from in-context, human-centred design research in LMICs[10]. Lighting parameters included measurements of illumination field (spot) diameter and illuminance (measured in lux, a measurement of visible light energy per area) as well as qualitative assessments of light uniformity and colour. Specific colour elements included red–blue colour differentiation, light washout (for example, loss of colour, diminished light intensity), reflection off tissue, and translucency using food models intended to mimic tissue (such as skin/soft tissue, mesentery, solid and hollow viscus). Mounting parameters included measurements of headlight weight and adjustment angle of the light source. Angle tilt was measured both above and below the line of sight (such as parallel to horizon) to accommodate multiple users, and to evaluate suitability for common pelvic (for example, hysterectomy) and perineal (childbirth) procedures. Battery parameters included measurements of run and recharge time, as well as visual indicators of charge levels during use, in between use and during battery recharging. In addition, standardized illuminance measurements were performed hourly for 10 h to determine illuminance with prolonged use. A full description of the testing parameters in each of these domains and specifications for measurement are provided in *Table S1* (supporting information).

A testing apparatus was assembled to maintain a fixed working distance of 40 cm between the light source and lux meter (*Fig. S1*, supporting information). Illuminance measurements for all headlights were captured in a dark room with no ambient lighting (0 lux) using a digital commercial lux meter (Dr.meter LX1330B©; Dr.meter, Union City, California, USA). An overall performance score with a maximum of 14 points was calculated for each headlight, with 1 point scored for each parameter that met the minimum specifications during testing. Domain scores in lighting (maximum 7), mounting (maximum 3) and battery (maximum 4) were also determined. However, given the importance of light intensity, any device selected for further field assessment must have satisfied an initial maximum illuminance of no less than 20 per cent below the predetermined specification (at least 28 000 lux). For reference, this illuminance target is close to industry standards for overhead surgical lighting, and well in excess of peak illuminance from mobile phone lights, which are often used as a primary lighting source in low-resource operating rooms; a common mobile phone light was measured to provide 140 lux at 40 cm.

Field assessment protocol and head-to-head comparison testing

Following device engineering testing, the highest-performing headlights were selected for field assessment. To ensure that a full range of key parameters was assessed, two additional medical headlights were added, which had similar illuminance, mounting and battery characteristics to those already selected, to explore the impact of variations in spot size, colour rendering and colour temperature index (approximation to the standard

of sunlight) as a means of evaluating the importance of these characteristics. These two additional headlights were produced by one of the medical manufacturers of a selected headlight, and were not available at the time of engineering testing. Because the additional performance capabilities related to features included in field testing only, they were added without the need for engineering testing first.

Surgeons were recruited from general, paediatric and vascular surgery, obstetrics and gynaecology, and orthopaedics to evaluate the headlights as part of a mixed-methods assessment both in their workplace settings and in a head-to-head comparison in an office setting. Surgeons were divided into two groups; each surgeon assessed one device per week on a rotating basis until all headlights had been evaluated in at least three rounds by different surgeons (*Table S2*, supporting information). Evaluations were performed in five hospitals located in Addis Ababa and a smaller teaching hospital in Western Ethiopia. The one technical headlight included in field assessment was distributed to both groups to allow more surgeons to undertake an evaluation.

Surgeons completed a logbook after each procedure in which they used a headlight to record feedback on lighting quality, mounting, comfort, battery run time, and both ambient and field lighting available in the operating room (*Appendix S2*, supporting information). Acceptability of each prespecified domain for each operation was converted to a percentage for summative grading. At the end of each weekly round during headlight exchange, the logbooks were reviewed and a rapid semistructured interview was undertaken (*Appendix S2*, supporting information). Surgeons were asked primarily about lighting quality, as well as storage and transport of the headlight, overall impressions, and positive and negative aspects of the device. They were asked to score the headlight on a scale from 1 to 7, where 7 represents outstanding. Devices were then collected and redistributed so that each headlight was evaluated by at least three surgeons.

After three rounds, participating surgeons met for head-to-head comparison testing of all headlights. Stations were set up with animal tissue samples selected to resemble tissue visualized in the operating room (*Appendix S2* and *Fig. S2*, supporting information). For headlights with adjustable spot sizes, surgeons first adjusted the spot illumination to their desired size and all subsequent measurements were recorded at a working distance of 40 cm. Each device was scored on domains of colour rendering, transillumination of tissue, light uniformity and reflectiveness (*Appendix S2*, supporting information). Surgeons also ranked the importance of lighting quality, mounting comfort, battery run time and price when choosing a headlight for their everyday practice. The list price for each headlight in US dollars was revealed; surgeons indicated which, if any, they would purchase for themselves with their own funds, and for a hospital if given a budget of US $1000 (€917).

Data analysis

The field assessment was approved by Core Human Factors institutional review board (IRB); no IRB approval was sought for engineering testing. Engineering testing data were recorded and stored electronically in AirTable (Formagrid, San Francisco, California, USA). Field assessment data were recorded in paper logbooks and interview forms, and qualitative feedback was recorded by transcribing direct quotes and field notes from conversations with surgeons. Field assessment data were entered into Microsoft Excel® version 16.28 (Microsoft, Redmond, Washington, USA), and analyses were undertaken using Excel® and Stata® version 15.1 (StataCorp, College Station, Texas, USA). Descriptive analyses were performed for case mix, operating room lighting and power outages. Mean scores for each headlight were calculated using ratings in logbooks, during interviews at the end of each assessment round, and during the final comparative testing session of all headlights. All scores were normalized to a total possible score of 5. Qualitative data from the interviews were summarized based on common themes identified through an inductive approach[12] and organized by headlight model. No inter-rater reliability assessment was undertaken as the themes were identified iteratively and agreed on mutually by three separate reviewers. For all testing, statistical significance was defined as a two-sided $\alpha < 0{\cdot}050$.

Results

Twenty-five device submissions were received from 13 unique global manufacturers of surgical and technical headlights following the request for expressions of interest. Seventeen headlights were excluded after screening for reported performance against predetermined specifications, price, ability to collaborate with Lifebox, financial viability and product track record. Six surgical and two technical headlights were included for full specification testing and final analysis.

Engineering testing

No headlight satisfied all minimum lighting, mounting and battery specifications; a summary of the individual testing

Table 1 Engineering testing of predetermined minimum specifications

	Lifebox proposed minimum specification	Headlights selected for engineering testing*								Additional headlights†	
		A (FM 1)	B	C (FM 2)	D (FM 3)	E	F	G	H	FM 4	FM 5
Lighting											
Minimum spot diameter (cm)‡	≤7	8·1§§	2·9	5·4	5·4	6·0	20§§	1·8	1·6	5·4	5·4
Maximum spot diameter (cm)‡	≥12	21·0	12·8	5·4§§	15·7	16·1	20	14·7	14·4	5·4§§	15·7
Maximum illuminance at 0 h (lux)§	≥35 000	48 900	4970§§	78 800	30 200§§	28 300§§	14 600§§	4030§§	1299§§	105 000	~39 000
Max illuminance after 3 h (lux)§	≥35 000	23 900§§	3870§§	59 500	28 800§§	27 500§§	2910§§	68§§	1303†	~80 000	~37 000
Minimum illuminance (lux)§	≤18 000	927	2670	33 600§§	13 290	14 530	396	4	315	~45 000§§	~17 000
Uniformity¶	Consistent illuminance	Y	Y	Y	Y	Y	N§§	Y	Y	Y	Y
Colour#	Tester rating ≥3	3	5	5	5	5	2§§	3	2§§	–	–
Colour temperature (×10^3 °K)#	4500–6500	5500	5000	4500	4500	4500	6500	5000	5000	6100	6100
Colour rendering index#	Ideally ≥90	–	90	90	90	90	–	93	–	75	75
Lighting score		5	5§§	5	5	5	2§§	5§§	4§§	–	–
Mounting											
Weight (g)	≤300	322§§	191	227	280	280	144	275	201	~227	~280
Angle adjustment**	≤ −60°	−50§§	−33§§	−71	−69	−69	−90	−51§§	−49§§	~ −70	~ −70
	≥5°	0§§	44	37	24	24	14	22	31	~30	~30
Mounting score		0	2	3	3	3	3	2	2	–	–
Battery											
Run time (h)††	≥3	10·0	16·4	6·1	5·0	11·9	10·0	10·0	13·9	≥5	≥5
Recharge time (h)	≤ Run time	7·5	8·5	5·2	5·2§§	6·0	3·2	3·5	2·3	5·2§§	5·2§§
Weight (g)	≤600	160	193	187	187	306	38	78	149	187	187
Charge-level indicator‡‡	During use	N§§	N§§	Y	Y	Y	N§§	N§§	N§§	Y	Y
	Between uses	N§§	N§§	Y	Y	Y	N§§	N§§	N§§	Y	Y
	During charge	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Battery score		3	3	4	3	4	3	3	3	–	–
Overall score		8	10	12	11	12	8	10	9	–	–

*Values measured by Lifebox during bench testing. †Similar to models C and D; values estimated from manufacturer specifications. ‡Measured at 40 cm working distance. §Measured at 40 cm working distance, 12 cm spot diameter, with exception of headlights C and F which had a fixed spot diameter of 5·4 and 20 cm respectively. ¶Measured at 40 cm working distance, 12 cm spot diameter, maximum illuminance. Uniformity determined by tester to determine if illuminance was consistent or variable throughout entire light field. #Measured at 40 cm working distance, 12 cm spot diameter, maximum illuminance, using models to simulate human tissue. Rating incorporated assessment of red–blue differentiation, reflection, transillumination and general representation of true colour items. Tester rating 1 (lowest quality) to 5 (highest quality); a score of at least 3 was considered acceptable. **Measured angle between vector parallel to headband and direction of light source. ††Measured at 40 cm working distance, 12 cm spot diameter, maximum illuminance. ‡‡Assigned score of 1 point if indicator present. §§Values did not meet Lifebox proposed minimum specifications. FM, field testing model; Y, yes; N, no.

parameters, as well as domain and overall scores is provided in *Table 1*. Of a possible 14 points, two headlights scored 12 (C, E), one scored 11 (D), two scored 10 (B, G), one scored 9 (H), and two scored 8 (A, F). Only four headlights (A, C, D and E) had an initial illumination intensity that met or exceeded 28 000 lux at a distance of 40 cm (at least 80 per cent of the predetermined specification of 35 000 lux).

In the lighting domain, five devices satisfied minimum spot diameter ranges (B, D, E, G and H). Two had fixed spot diameters (C, 5·4 cm; F, 20 cm) and one (A) failed to meet the minimum spot diameter specification. Only headlights A and C met the illuminance parameters of greater than 35 000 lux in the first 3 h during testing. Headlights D and E showed minimum loss of illuminance during prolonged use, despite being below the 35 000-lux illuminance threshold (*Fig. S3*, supporting information). Headlight F had variable illuminance uniformity, specifically with more than 25 per cent variability in lux meter readings throughout the light field; headlight A was rated as uniform as its central bright spot was large and uniform, despite having a second dimmer light ring halo around the outside. On qualitative colour assessment, headlights A and G were rated 3 of 5 owing to light reflection and poor red–blue colour differentiation respectively,

Table 2 Overall procedure characteristics and surgical lighting available

	Model 1	Model 2	Model 3	Model 4	Model 5	Total	*P*†
Spot size adjustability	Adjustable	Fixed	Adjustable	Fixed	Adjustable		
No. of operations	30	17	12	13	14	86	
Surgical specialty							< 0·001
General surgery	14 (47)	4 (24)	6 (50)	9 (69)	3 (21)	36 (42)	
Orthopaedics	4 (13)	5 (29)	0 (0)	0 (0)	5 (36)	14 (16)	
Paediatrics	10 (33)	8 (47)	0 (0)	0 (0)	6 (43)	24 (28)	
Obstetrics	2 (7)	0 (0)	6 (50)	4 (31)	0 (0)	12 (14)	
Duration of operation (min)*	60 (40–90)	90 (53–120)	50 (30–90)	35 (30–88)	100 (70–120)	60 (40–105)	0·047‡
Time of day							0·013
Day	30 (100)	15 (88)	9 (75)	13 (100)	14 (100)	81 (94)	
Night	0 (0)	2 (12)	3 (25)	0 (0)	0 (0)	5 (6)	
Available OR lighting							
Overhead OR light	29 (97)	17 (100)	10 (83)	11 (85)	14 (100)	81 (94)	0·140
Overhead ceiling light	8 (27)	9 (53)	1 (8)	7 (54)	3 (21)	28 (33)	0·035
Window/ambient light only	0 (0)	0 (0)	0 (0)	0 (0)	1 (7)	1 (1)	0·450
No light available	0 (0)	0 (0)	0 (0)	2 (15)	0 (0)	2 (2)	0·062
Power interruptions							
Power outage during procedure	2 (7)	1 (6)	0 (0)	2 (17)	3 (21)	8 (9)	0·310
Generator turned on during procedure	2 (7)	1 (6)	0 (0)	0 (0)	2 (14)	5 (6)	0·600

Values in parentheses are percentages unless indicated otherwise. *values are median (i.q.r.). OR, operating room. Pearson's χ^2 or Fisher's exact test, except ‡Kruskal–Wallis test.

whereas headlights F and H were rated 2 of 5 given poor illumination and perceived light washout, noted especially in darker-coloured simulated models; all others were rated 5 of 5.

In the mounting domain, only one headlight exceeded the minimum specification for maximum weight (headlight A, 322 g). In regard to angle adjustment, headlight A was also the only model that did not provide a light source adjustable to 5° above the user's line of sight. All batteries had a continuous run time of at least 3 h; only the battery powering headlight D had a recharge time longer than the usage run time (5·2 *versus* 5·0 h). Three headlights (C, D and E) had a charge-level indicator during and between headlight use.

Ultimately, four headlights (B, F, G and H) were excluded from field assessment given their significantly lower initial maximum illuminance (less than 28 000 lux). Of the remaining ones (A, C, D and E), headlight E was identical to headlight D but with a more expensive battery. Therefore, three headlights were selected for further field assessment in Ethiopia: headlights A, C and D. As noted above, two additional devices were added to further explore the importance to surgeons of spot size, colour rendering and colour temperature. Of the five devices identified for field assessment, four were marketed as medical headlights whereas one was a technical, non-medical headlight.

Field assessment and head-to-head comparison of devices

Headlights were evaluated in a total of 86 operations over the course of 6 weeks (*Table 2*). Although the original intention had been to complete this work over 3 weeks, owing to civil unrest and a declared state of emergency[13,14], the participants were asked to shelter in place and not undertake any movements on behalf of the project; thus, the second round of assessment was extended to a total of 4 weeks until the state of emergency had been lifted and travel restrictions eased, allowing the exchange of headlights. Because of their schedules, some surgeons were not able to assess devices in every round. In addition, only seven of the eight surgeons participated in the final comparison testing owing to scheduling conflicts.

The volume of experience with each headlight was similar in the two groups, with 12–17 operations undertaken for each device (model 1 was evaluated in both groups). Overall, 36 procedures (42 per cent) were from general surgery, 24 (28 per cent) from paediatric surgery, 14 (16 per cent) from orthopaedics and 12 (14 per cent) from obstetrics (*Table 2*). Mean(s.d.) operating time was 75(42) min and 94 per cent of operations were performed during the day. An overhead operating room light was available during 81 procedures (94 per cent). A power outage occurred during eight operations (9 per cent), and in five of these the

Table 3 Evaluations of surgical headlights from logbooks, interviews, and final comparison testing

	Model 1	Model 2	Model 3	Model 4	Model 5
No. of operations*	30	17	12	13	14
Logbooks: intraoperative headlight evaluations					
% considered appropriate/adequate					
Illumination	70	6	92	46	21
Brightness	73	47	91	85	93
Spot size	83	0	92	31	0
Tilt	100	82	100	100	100
Secure fit	93	53	100	62	93
Battery run time	97	82	83	92	93
Mean % adequate	86	59	93	69	67
Mean score†	4·3	2·9	4·7	3·5	3·3
Interviews: summative headlight ratings					
Would use routinely	5 of 6	1 of 3	2 of 3	1 of 3	3 of 3
Overall rating†	3·8	2·9	3·8	3·1	3·6
Final comparison testing: head-to-head lighting quality comparison† (7 surgeons)					
Mean(s.d.) preferred spot diameter (cm)	9·6(1·6)	–	15·9(0·9)	–	13·8(2·1)
Colour	4·4	4·1	4·7	4·3	4·3
Transillumination	3·6	3·4	4·0	3·7	4·5
Uniformity	3·1	3·4	4·3	3·9	4·7
Mean score†	3·7	3·7	4·3	4·0	4·5
Final comparison testing: rating of all surgical headlights† (7 surgeons)					
Lighting	2·4	1·0	2·7	1·6	3·4
Mounting	2·0	1·4	2·7	2·1	3·2
Battery	2·8	2·7	3·0	3·3	3·2
Price	3·7	2·0	1·4	1·7	1·5
Overall rating	2·7	0·7	2·5	1·5	3·2
Mean score†	2·7	1·6	2·5	2·1	2·9
Final comparison testing: pricing evaluation (6 surgeons)					
Would purchase independently	4 of 6	0 of 6	1 of 6	0 of 6	2 of 6
Would purchase with hospital budget	3 of 6	0 of 6	2 of 6	1 of 6	4 of 6

*Model 1 evaluated by six surgeons; all others by three surgeons each. †Score from 1 (poor) to 5 (excellent).

hospital generator was turned on. On rare occasions where headlights were not used, surgeons reported that this was during short, cutaneous procedures where they felt additional visibility would not be needed, or in endoscopic procedures where a headlight would not be helpful.

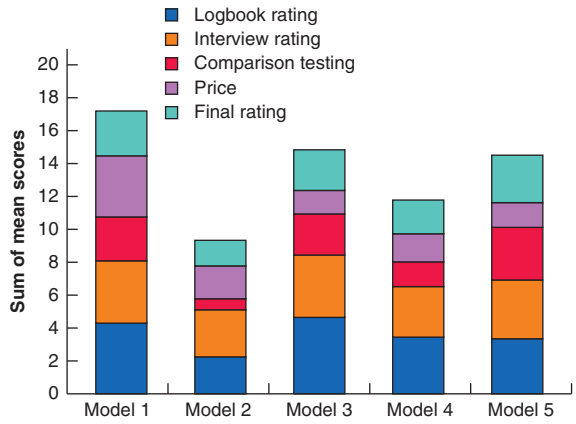
Individual operating room logbooks

Surgeons reported adequate tilt angle in 82–100 per cent of operations and adequate battery run time in 82–97 per cent of procedures across all headlight models. However, lighting adequacy (6–92 per cent), spot size (0–92 per cent) and security of fit (53–100 per cent) ratings were highly variable across the five headlights (*Table 3*). The spot size was rated as too small in all 17 operations for model 2, and nine of 13 for model 4. Although the spot size range was the same in models 3 and 5, surgeons rated the spot size for model 5 as too small across all 14 operations, whereas model 3 was deemed to have an appropriate spot size in 11 of 12 procedures. Normalization of ratings to a total score of 5 showed that the top-ranking headlights were model 3 (4·7) and model 1 (4·3). Full details are provided in (*Table S3*, supporting information).

End-of-round feedback interviews

A total of 18 interviews were conducted with surgeon participants, six at the end of each of three rounds. In all interviews, surgeons reported that headband adjustability and battery charging were adequate. When asked if they would use each device routinely in operations, the highest response levels were for model 1 (5 of 6) and model 5 (3 of 3) (*Table 3*). After rating overall satisfaction with headlights, the highest-scoring devices were models 1 and 3 (each 3·8 of 5). Key themes from the qualitative feedback focused on lighting intensity, spot size adjustability, as well

Fig. 2 Aggregate performance of surgical headlights in logbooks, interviews, final comparison testing and pricing evaluations



as size and transportability of headlight packaging (*Table S4*, supporting information).

In interviews, surgeons had a variety of issues with the portability of models 2–5 which do not collapse down and came with larger carrying cases. The issues ranged from inability to fit in a backpack to the worry about perceptions of theft. In contrast, model 1 received good reviews regarding portability, convenience and packaging (*Tables S4* and *S5*, supporting information).

Final comparison testing

The importance of lighting quality, mounting comfort, battery run time and price were rated fairly equally, with price being the highest consideration; portability was identified as an additional driving factor in decision-making (*Table S6*, supporting information). In the final head-to-head comparison, the highest scoring devices were models 1, 3 and 5 on lighting elements including colour rendering, transillumination and uniformity. Participants ranked model 3 as best in terms of colour rendering (colour rendering index score 90) and model 1 as second best (*Table 3*). When surgeons set the adjustable headlights at their ideal spot diameter, measured at a distance of 40 cm, the mean(s.d.) preferred spot diameter was 9·6(1·6) cm for model 1, 15·9(0·9) cm for model 3 and 13·8(2·1) cm for model 5. After the prices had been revealed, four of six surgeons indicated that they would purchase model 1 with their own money, and two would purchase model 5 (relative prices of each device are as follows: model 1, x; models 2 and 4, 3·5x; models 3 and 5, 4·8x). Given a hospital budget of US $1000 (€917), four surgeons indicated that they would purchase model 5, and three would also purchase model 1. No surgeon would purchase model 2 or 4 with their own funds, citing the small, overly bright, fixed spot diameter (5·4 cm measured at a distance of 40 cm).

Selection of final headlight model

Overall, models 1, 3 and 5 performed best in intraoperative ratings, lighting quality comparison and final rating of headlights. Features of these three models were somewhat similar in terms of adjustable spot size, moderate brightness illuminance and highly rated lighting quality. Summative rankings for each headlight model are shown in *Fig. 2*. Overall, headlight model 1 was rated as the preferred headlight when considering both quality parameters and price.

Discussion

Of five headlights that underwent field assessment in Ethiopia, three outperformed the others with respect to function, whereas a single one was preferred owing to its affordability and ease of transport. Based on human-centred design principles, the model 1 headlight has been selected by Lifebox for further modification, distribution and ongoing assessment. This was principally due to cost considerations, as expense is a major barrier to surgeons' access to surgical headlights in LMICs.

Currently available medical and technical headlight manufacturer specifications for illuminance and battery parameters are not standardized with respect to specific working distances, spot diameter or intensity levels for battery testing. This lack of standardization makes it challenging for surgeons to identify a device appropriate for their needs, particularly in LMICs where there are substantial costs and commercial barriers to testing headlights in the field. Although all headlights tested met the majority of predetermined minimum specifications, no headlight satisfied all requirements. Based on field assessments of headlight models by Ethiopian surgeons, the most highly rated headlights succeeded owing to brightness, battery run time and ease of transport. In contrast, headlights rated lowest were rated as such due to small spot size, large or cumbersome packaging, and dim or non-uniform lighting. The portability of headlights was not rated formally, but in qualitative assessment a small portable size seemed to influence the surgeons' choice of purchase. These features have informed design and engineering modifications, including a few improvement opportunities such as strap material and configuration, portable casing and improved illumination intensity. Ultimately, Ethiopian surgeons were extremely sensitive to price, but still demanded a high-performing headlight.

Surgical lighting is clearly a critical issue, and a high-quality device could play a major role in improving surgical safety. The process described here confirmed previously published findings that, even in larger tertiary hospitals in a capital city, lighting is an ongoing patient safety issue[3]; during the short evaluation period a power outage occurred during nearly 10 per cent procedures, and a back-up generator was available in only five of eight instances. In Malawi, a study[15] of the energy supply in all health facilities demonstrated that 63 per cent had either interrupted power supply without back-up energy sources, or no electricity at all. Although that study did not measure operating room lighting, these facilities were able to light delivery rooms only 18 per cent of the time. Similarly, various studies[16–20] in Senegal, Rwanda, Uganda and Sierra Leone have identified major gaps in access to reliable grid electricity and generator power. Perhaps more importantly, a qualitative study[21] in Rwanda highlighted that resource variability, rather than a simple and consistent lack of resources, is a major challenge to providing safe surgical care. A widely available surgical headlight could also extend beyond operating room use to support safe night-time clinical services and safe childbirth in contexts where grid electricity is intermittent or not available.

There are several barriers limiting the availability of surgical headlights in LMICs. First, none have been engineered specifically for this environment, which is not unusual as manufacturers do not always consult the end-user for input in the design process[22], citing barriers of ethical approval or a perspective that high-level administrators may provide adequate feedback during the design process. A further lack of uniformity in specifications listed by the manufacturer makes devices difficult to compare, and the authors found the quality of lighting to be extremely variable.

Current surgical headlights are costly; most retail at well above US $1200 (€1100) and are therefore out of reach of the average LMIC surgeon. There are few opportunities in LMICs for surgeons to evaluate, test and ultimately obtain headlights fit for their local needs, including those with extended battery life, water and dust resistance, and improved durability. In non-resource-limited environments, evaluation of new or existing surgical headlight technology often occurs by exposure to manufacturers via industry representatives or during conferences and conventions. These settings allow surgeons to experience the performance of multiple headlights directly, receive education related to lighting, discuss financial support, negotiate costs and arrange maintenance[23]. The majority of conferences with industry presence occur in high-resource settings; even global health meetings near or in LMICs rarely have representatives from global or in-country headlight manufacturers[24].

Finally, there is no current market mechanism or incentive for delivering these devices to those most likely to benefit. Supply chains in LMICs, particularly for medical devices, can be complicated, expensive or non-existent, with frequent delays in customs processing and high import fees. For all of these reasons, it is not surprising that only about 9 per cent of surgical providers in LMICs report regular access to, or use of, surgical headlights[3]. Furthermore, donation of medical devices in general is fraught with difficulty; many medical devices arrive through donation, and anywhere between 40 and 96 per cent of these devices have been reported to be non-operational[25]. Developing a surgical headlight that is available for use in a myriad of clinical situations will reduce the variability in safe lighting availability when grid or generator power is not available. Although efforts to improve infrastructure for surgical lighting (such as electricity and generator support) may represent the most durable solution, there is limited funding dedicated to infrastructure improvements in LMICs. Gutnik and colleagues[26] found that, of over US $105 million (€96 million) from charitable foundations, less than 2 per cent was allocated for infrastructure improvements. It is essential to consider the possible unintended impact of any intervention in the global surgery realm, and also to identify and address the true root of the problem in surgical safety. The surgical lighting gap is truly multifactorial. Mains power is a definite contributor, and lack of access to surgical headlights a prevalent issue, but supply chains for appropriate replacement bulbs for built-in examination lights and biomedical engineering capacity to repair faulty overhead lights are all lacking in many LMIC operating rooms. The Surgical Headlight Project is a concurrent effort not meant to replace, disincentivize or deter the larger infrastructure changes that need to take place in many nations. The project aims to augment lighting for safer surgery, as surgical headlights are often useful even when overhead lighting is in place, and also as a stop gap to improve surgical safety in the short term during transient power outages, as the changes needed in infrastructure improve over the longer term.

This study has several limitations. Specification testing was undertaken only on headlights from manufacturers who provided an expression of interest, and testing was limited to devices with a proposed wholesale price of no more than US $600 (€550). Although this may represent a narrow sample, extensive market research was performed for both surgical and technical headlights within screening criteria determined *a priori*. In addition, testing that would potentially damage the loaned headlights

(such as performance in extremes of temperature or exposure to splashes) was not performed. The field assessment was conducted primarily in a capital city, Addis Ababa, with residency-trained surgeons, and may have failed to capture additional feedback from non-physician surgical providers. Most of the operations performed with a headlight included overhead surgical lighting and took place during the day; further evaluations with a wider variety of surgical providers who have less reliable surgical lighting and operate at night would be useful. Strengths of the study include its mixed methodology, human-centred design, and capture of iterative feedback on lighting features across time and work environments. In human-centred design, a number of barriers can impede the ability of the end-user to contribute to medical device design, including communication between medical and engineering disciplines, and assurance of shared ownership of the design space[27]. An attempt was made to mitigate these barriers by obtaining open-ended qualitative feedback from medical providers, using medical professionals and trainees to perform interviews and data collection, and obtaining additional information on ergonomics and comfort[28].

Through the Lifebox Surgical Headlight Project, several headlights, representing a range of performance across lighting colour and intensity, spot size, mounting and battery domains, were identified for end-user field assessment and selection using human-centred design principles[29]. By employing specifications, engineering testing and field assessment in a cohort of Ethiopian surgeons, an appropriate, fit-for-purpose, affordable surgical headlight has been selected, but further modification-based user feedback is required to meet surgeon preferences and needs. Future testing of this first-generation model will include user testing over a period of 6–18 months with several hundred surgeons to obtain further feedback on use in the workplace, function and durability, and necessary design changes will be made in collaboration with the manufacturer. This process will culminate in the production and distribution of a surgical headlight specifically designed for low-resource, austere environments, and will contribute to patient safety in surgery.

Acknowledgements

N.S. and N.P. are joint first authors of this article. The authors thank the Surgical Society of Ethiopia and all participating surgeons for their time and contribution to this work, and to the College of Surgeons of East, Central and Southern Africa for its collaboration in the Lifebox Headlight Testing Project. N.S. is supported by National Institutes of Health (NIH) grant T32 DK7573-29 and N.P. by NIH grant T32 DK007754-18. T.G.W. was supported by an innovation seed grant from the Stanford University Department of Surgery. The Lifebox Surgical Headlight Project is partially funded by a grant from the UBS Optimus Foundation. No funder was involved in the analysis or decision to submit this work for publication. N.S. is a current and J.A.F. a former Lifebox Fellow; E.W.J. is the Principal of Spark Health Design, a design consulting firm; K.F. is the Head of Programmes at Lifebox; S.B. is the Country Programme Manager of Lifebox Ethiopia; T.N.M. is the Clinical Lead of Lifebox Ethiopia; and T.G.W. serves as the Clinical Advisor for Lifebox.

Disclosure: The authors declare no other conflict of interest.

References

1 Weiser TG, Haynes AB, Molina G, Lipsitz SR, Esquivel MM, Uribe-Leitz T *et al.* Size and distribution of the global volume of surgery in 2012. *Bull World Health Organ* 2016; **94**: 201–209F.

2 WHO, World Bank. *Access to Modern Energy Services For Health Facilities in Resource-Constrained Settings: a Review of Status, Significance, Challenges and Measurement*; 2014. https://apps.who.int/iris/handle/10665/156847 [accessed 14 Sepetember 2019].

3 Forrester JA, Boyd NJ, Fitzgerald JEF, Wilson IH, Bekele A, Weiser TG. Impact of surgical lighting on intraoperative safety in low-resource settings: a cross-sectional survey of surgical providers. *World J Surg* 2017; **41**: 3055–3065.

4 Adair-Rohani H, Zukor K, Bonjour S, Wilburn S, Kuesel AC, Hebert R *et al.* Limited electricity access in health facilities of sub-Saharan Africa: a systematic review of data on electricity access, sources, and reliability. *Glob Health Sci Pract* 2013; **1**: 249–261.

5 Vargas J, Mayegga E, Nuwas E, Ellegala DB, Kucia EJ, Nicholas J. Brain surgery in the bush: adapting techniques and technology to fit the developing world. *World Neurosurg* 2013; **80**: e91–e94.

6 Forrester JA, Torgeson K, Weiser TG. Minimum specifications for a Lifebox surgical headlight for resource-constrained settings. *JAMA Surg* 2019; **154**: 80–82.

7 Privitera MB, Evans M, Southee D. Human factors in the design of medical devices – approaches to meeting international standards in the European Union and USA. *Appl Ergon* 2017; **59**(Pt A): 251–263.

8 North R, Pospisil C, Clukey RJ, Parkin CG. Impact of human factors testing on medical device design: validation of an automated CGM sensor applicator. *J Diabetes Sci Technol* 2019; **13**: 949–953.

9 IEC, ISO. *IEC 62366-1:2015(en,fr). Medical Devices – Part 1: Application of Usability Engineering to Medical Devices/ Dispositifs Médicaux – Partie 1: Application de l'Ingénierie de l'Aptitude à l'Utilisation aux Dispositifs Médicaux*; 2015. https://

www.iso.org/obp/ui/#iso:std:iec:62366:-1:ed-1:v1:en,fr [accessed 17 January 2020].
10 IDEO.org. *The Field Guide to Human Centered Design* (1st edn). IDEO.org: Canada, 2015.
11 Lifebox Foundation. *Lifebox Light: Request for Expressions of Interest*. https://www.lifebox.org/wp-content/uploads/2018/12/Lifebox-Light-request-for-EoIs-Dec-2018.pdf [accessed 25 March 2020].
12 Thomas DR. A general inductive approach for analyzing qualitative evaluation data. *Am J Eval* 2016; **27**: 237–246.
13 Lewis D, Endeshaw D; Reuters. *'Dozens' Killed in Foiled Ethiopia Coup Attempt: Regional Government*; 2019. https://www.reuters.com/article/us-ethiopia-security-idUSKCN1TR1AL [accessed 29 January 2020]
14 Parkinson J; Wall Street Journal. *Ethiopia, in the Midst of a Political Crisis, Foils Coup Attempt*; 2019. https://www.wsj.com/articles/ethiopia-in-the-midst-of-a-political-crisis-foils-coup-attempt-11561287955 [accessed 29 January 2020].
15 Suhlrie L, Bartram J, Burns J, Joca L, Tomaro J, Rehfuess E. The role of energy in health facilities: a conceptual framework and complementary data assessment in Malawi. *PLoS One* 2018; **13**: e0200261.
16 Ouedraogo NS, Schimanski C. Energy poverty in healthcare facilities: a 'silent barrier' to improved healthcare in sub-Saharan Africa. *J Public Health Policy* 2018; **39**: 358–371.
17 Petroze RT, Nzayisenga A, Rusanganwa V, Ntakiyiruta G, Calland JF. Comprehensive national analysis of emergency and essential surgical capacity in Rwanda. *Br J Surg* 2012; **99**: 436–443.
18 Linden AF, Sekidde FS, Galukande M, Knowlton LM, Chackungal S, McQueen KAK. Challenges of surgery in developing countries: a survey of surgical and anesthesia capacity in Uganda's public hospitals. *World J Surg* 2012; **36**: 1056–1065.
19 Kingham TP, Kamara TB, Cherian MN, Gosselin RA, Simkins M, Meissner C *et al.* Quantifying surgical capacity in Sierra Leone: a guide for improving surgical care. *Arch Surg* 2009; **144**: 122–127.
20 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.
21 Scott JW, Lin Y, Ntakiyiruta G, Mutabazi ZA, Davis WA, Morris MA *et al.* Contextual challenges to safe surgery in a resource-limited setting: a multicenter, multiprofessional qualitative study. *Ann Surg* 2018; **267**: 461–467.
22 Money AG, Barnett J, Kuljis J, Craven MP, Martin JL, Young T. The role of the user within the medical device design and development process: medical device manufacturers' perspectives. *BMC Med Inform Decis Mak* 2011; **11**: 15.
23 Dua A, Rothenberg K, Lee JT. The role of industry in vascular surgery trainee education. *Semin Vasc Surg* 2019; **32**: 30–32.
24 American College of Surgeons. *Future Clinical Congresses*. https://www.facs.org/education/clincon/future-clinical-congresses [accessed 6 February 2020].
25 Perry L, Malkin R. Effectiveness of medical equipment donations to improve health systems: how much medical equipment is broken in the developing world? *Med Biol Eng Comput* 2011; **49**: 719–722.
26 Gutnik L, Dieleman J, Dare AJ, Ramos MS, Riviello R, Meara JG *et al.* Funding allocation to surgery in low and middle-income countries: a retrospective analysis of contributions from the USA. *BMJ Open* 2015; **5**: e008780.
27 Vincent CJ, Li Y, Blandford A. Integration of human factors and ergonomics during medical device design and development: it's all about communication. *Appl Ergon* 2014; **45**: 413–419.
28 Martin JL, Norris BJ, Murphy E, Crowe JA. Medical device development: the challenge for ergonomics. *Appl Ergon* 2008; **39**: 271–283.
29 Fisher MJ, Johansen E. Human-centered design for medical devices and diagnostics in global health. *Global Health Innovation* 2020; **3**. https://doi.org/10.15641/ghi.v3i1.762

Supporting information

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