Developing a Core Measurement Set for perioperative care safety in Europe within the SAFEST project: An international Delphi consensus study

Master's Degree in Public Health

José Pedro Dinis Teixeira

January 2024

ESCOLA NACIONAL DE SAÚDE PÚBLICA

Developing a Core Measurement Set for perioperative care safety in Europe within the SAFEST project: An international Delphi consensus study

Dissertation submitted to fulfil the requirements for obtaining the degree of Master of Public Health, under the scientific orientation of Ana Beatriz Nunes and Paulo Sousa.

January 2024

Abstract

A Core Measure Set (CMS) is an agreed standardized group of measures that should be measured and reported in future research or practice within a clinical area or a specific condition. This study undertook the development of a CMS for Patient Safety through a two-round, web-based Delphi consensus approach, in the context of the “Improving quality and patient SAFEty in surgical care through STandardisation and harmonisation of perioperative care in Europe” (SAFEST) project - a collaborative, patient-centered and evidence-based European Union-funded project that aims to generate action-oriented evidence in perioperative care. We developed an Initial List of Indicators via an umbrella review following the deployment of an e-Delphi method with an inclusive panel of experts to prioritize indicators towards a consensualized Final List of Indicators (FLI). All indicators were assessed for both importance and feasibility. After the two rounds of the e-Delphi consensus method we observed 13 preoperative indicators (~40.6% of the initial number), 24 intraoperative indicators (~66.7%), 25 postoperative indicators (~20.3%) and 23 mixed period indicators (~41.1%) met consensus criteria for both importance and feasibility. Higher scores were detected in importance ratings compared to feasibility across all groups of indicators. Importantly, numeric averages regarding pain-related indicators differed in the assessment of patients when compared to that of HCPs. This work not only informs future SAFEST iterations but also sets a precedent for research into valid, patient-centered, and action-oriented perioperative safety indicators.

Keywords

Patient safety, core measurement set, Delphi study, perioperative care, public health

Index

1. Introduction

1.1. Global context

The World Health Organization (WHO) defines patient safety as a framework of organized activities towards cultures, processes, procedures, behaviors, technologies and environments in health care that consistently and sustainably lower risks, reduce the occurrence of avoidable harm, make errors less likely and reduce the impact of harm when it does occur (1).

Advancing patient safety has become a paramount concern in healthcare research and practice. Since the notable publication of the "To Err is Human" report by the Institute of Medicine (IOM) in 2000 (2), much attention has been brought to the general public as well as governmental agencies on the central importance of patient safety for the delivery of quality healthcare services.

In 2015, all 193 member states of the United Nations (UN) committed to engage in efforts to achieve Universal Health Coverage (UHC) (3). By definition, this means that all individuals and respective communities should receive adequate health services without suffering financial hardship.

However, although UHC comprises the full spectrum of essential, good quality health services, these should not be achieved at the expense of the foundational bioethics precept first do no harm (*primum non nocere*), from which derives the principle of non-maleficence in Medicine (4). Therefore, patient safety is considered paramount to the achievement of UHC.

1.2. Prevalence and consequences of unsafe care

A significant volume of evidence points toward a non-negligible prevalence of unsafe care in different healthcare delivery settings. Approximately 10% of patients receiving in-hospital care are estimated to experience adverse events, with a relevant proportion of these incidents being avoidable (5,6). A recent work by Bates et al. showed that notwithstanding the global efforts to tackle this issue in recent years, it remains rather prevalent and, as such, extremely relevant. Analysing a random sample of 2809 hospital admissions, they identified at least one adverse event in 23,6% (7).

Based on an analytic modelling approach of observational studies, adverse events are reported to amount to 42.7 million annually worldwide, and may account for the loss of 23 million disability-adjusted life years (DALYs) per year (8).

While hospital data is more thoroughly explored in the literature, patient safety is critical across all levels of care.

According to an Organization for Economic Co-Operation and Development (OECD) report, safety lapses occur between 1 and 24 times in every 100 primary care consultations and even though harm in primary care might be less visible compared to harm related to hospital-based interventions the total impact is not inferior (9).

A pivotal European study published in 2012 including 498 hospitals from 28 countries found that the mortality rate for patients undergoing surgery was higher than previously thought (10). Further evidence points towards a proportion of severe seriousness among in-hospital adverse events that could amount to over 30% (7).

Some approaches tailored to tackle surgery-specific impacts of harmful care have already been implemented. The creation and enforcement of a checklist based on the WHO Surgical Checklist effectively decreased mortality in a group of patients undergoing non-day case surgery in a Dutch university hospital (11). Another impacted area by the WHO Surgical Checklist was anaesthesia-related mortality. It has reportedly dropped from 357 to 34 deaths per million in the United States over the last three decades (12).
In addition to directly affecting the lives of patients and their carers, unsafe care also bears a substantial economic impact. Based on OECD insights publicly reported, 1 in 10 US dollars spent on health care of high-income countries was spent in managing the consequences of safety errors (13).

Taking into consideration the significant magnitude and burden of the adverse events, particularly those related to surgery, in terms of mortality and morbidity, its' undeniable socioeconomic relevance as well as how enforceable it is, Patient Safety must be seen as a global public health concern.

1.3. Core Measurement Sets and the Delphi method

A Core Measure Set (CMS) is an agreed standardized group of measures that should be measured and reported in all future research of practice for a specific condition or clinical area (14). These sets do not imply that outcomes in a particular trial should be restricted to those in the CMS. Rather, there is an expectation that the core outcomes will always be collected and reported, and that researchers will continue to explore other outcomes (15). The use of CMS might also bear an important impact on benchmarking for healthcare delivery services (16).

Multiple publications have reported the use of the Delphi method as a key pillar to the development of a CMS in different areas (17–20).

The Delphi method is a systematic method used for structuring the process of communication within a group of individuals. The method is reported to have been first developed in the 1950s by the RAND Corporation as a way to obtain the collective judgment of experts on complex issues related to military policy (21), and is often employed to gather and synthesize the opinions of experts in a particular field.

Since its inception, the Delphi method has been widely used in a variety of fields, including management, engineering, social science, and public health. It is particularly useful in situations where there is a need to make decisions or forecast events based on incomplete or uncertain information (22).

In the realm of medical sciences, Delphi methods might be employed as an alternative when conducting experimental or quasi-experimental studies are considered to be impractical due to ethical, pragmatic or economic constraints.

Firstly conceptualized in a report by the Canadian Task Force on the Periodic Health Examination (23), levels of evidence are ranked according to characteristics such as the propension to systemic bias (24). According to this proposed hierarchy of scientific evidence, expert consensus (of which the Delphi technique's objective is an example of) is classically considered to be at the bottom of the pyramid. However, it might be deceitful to consider that expert consensus is inherently inferior. The strength and quality of an expert consensus rely heavily on the evidence used by the experts to reach that consensus and this might include higher level pieces of evidence, such as systematic reviews, meta-analyses, as well as personal professional experience (25). Additionally, the Delphi rounds themselves may build upon inputs from other studies performed earlier, such as systematic reviews (26,27).

Importantly, studies that rely on the Delphi method can function as tools to further empower patients and patients' representatives by bringing them to the forefront of evidence generation. This has been seen in a wide range of areas, from the assessment of the impact of therapies to the development of undergraduate medical curricula or health literacy for young cancer patients (28–30).

The Delphi method operates on the premise that forecasts or decisions derived from a formal group of individuals are more precise and reliable compared to those from informal groups (31). The method consists of several rounds of anonymous questionnaire surveys, in which experts are asked to provide their inputs on a specific topic or question. The responses are then collated and analyzed, and feedback is provided to the experts for review. This process is usually repeated until a consensus is reached or until the level of agreement among the experts reaches a satisfactory level. The definition of what constitutes a satisfactory level of agreement remains, however, up for debate in the literature (32,33). An optimal approach to

consensus in the Delphi method might come from formally defining consensus criteria a priori, while not assuming it to be an automatic outcome at the conclusion of the method (34).

As any other decision-making procedure, this method presents its strengths and weaknesses. On the advantages side, the Delphi method allows for the input from multiple experts that might have diverse academic and professional backgrounds, experiences, and viewpoints on the topics being studied. By considering a wide range of perspectives and different insights, this can help ensure a more comprehensive and inclusive approach to decision-making that can be particularly useful when dealing with complex problems (35).

Furthermore, the iterative process of the Delphi method, which includes multiple rounds of feedback and possible exchange of ideas, can facilitate the generation of consensus among experts. By identifying areas of agreement and disagreement, the method allows for efforts to reconcile differences and work towards a shared decision or outcome.

1.4. The SAFEST project

“Improving quality and patient SAFEty in surgical care through STandardisation and harmonisation of perioperative care in Europe” (SAFEST) is a collaborative, patient-centered and evidence-based European Union (EU)-funded project that aims to generate evidence within this field of knowledge.

SAFEST assembled a 10-member consortium that comprises research institutions, hospitals, as well as health policy organizations and a healthcare management and analytics company. It is led by Fundación Avedis Donabedian para la mejora de la calidad asistencial (FAD). The other nine members of the consortium are: Netherlands Institute for Health Services Research (NIVEL), NOVA University of Lisbon (NUL), Radboud university medical center (Radboudumc), Sistema Español de Notificación en Seguridad en Anestesia y Reanimación (SENSAR), University of Tartu (UT), OptiMedis AG (OM), European Hospital and Healthcare Federation (HOPE), European society of Anaesthesiology and Intensive Care (ESAIC), and Spojená akreditační komise (SAK).

The main objective of the SAFEST project is to improve the adherence to evidence-based standardized patient safety practices in perioperative care by 15% and reduce the frequency of surgical complications by 8% after 18 months of the intervention (36). Across an expected period of four years of the project several activities shall add up towards its primary goal.

As an intervention-oriented project with relevant expected impact in the quality and safety of provided perisurgical care, the scope of SAFEST is largely dependent on the measurement and evaluation of indicators.

Specifically, the development of CMS serves two main objectives. Initially, it is expected to be used in the SAFEST study to oversee and assess safety outcomes in ten European hospitals where the SAFEST recommendations will be implemented. Additionally, it will play a role in upcoming European studies and initiatives focused on surgical safety. This shall be achieved by promoting standardized outcome reporting in perioperative patient safety. Furthermore, the CMS will facilitate the comparison of data from various studies.

As part of this project, the development of a CMS for achieving an EU-wide consensus on relevant and feasible core measures to assess patient safety in perioperative care in surgical adult patients was planned.

The objective of this study is to decisively contribute to a consensus among experts (both professionals and people with lived experience), ultimately resulting in the formulation of a consensualized list of indicators which might work as a stepping-stone for assessing perisurgical care safety and quality in hospitals.

2. Methods

2.1. General approach and selection of experts

An integrated approach was used to pursue the development of a CMS for Patient Safety in Perioperative Care. The methodology of this qualitative study based on a web based Delphi survey (e-Delphi) comprised three steps. Initially, a comprehensive list of indicators for patient safety in perioperative care among surgical adult patients was created. This was mainly accomplished through an umbrella review of indicators conducted by SAFEST's Work Package (WP) 6, which examined existing indicators within its scope, while also drawing data from WP2 and WP4 reviews. Secondly, a group of healthcare professionals and patients' representatives was assembled in order to prioritize these indicators using a two-round e-Delphi method.

The research team aimed to ensure a diverse panel, with heterogenous expertise and backgrounds as well as balanced sociodemographic characteristics like gender.

The panel incorporated members from the Scientific Advisory Group (SAG) who were previously selected within the workstream of WP2. Experts from key stakeholder groups including healthcare professionals, patients and patients' representatives, regulatory agency representatives, policymakers, private sector representatives, guideline developers, CMS methodologists/researchers, and governmental agency representatives were recruited during the first six months of the SAFEST project.

All experts were selected based on their experience, expertise in patient safety, perioperative care, or outcomes research. Additionally, experts needed to be adults and capable of completing the online surveys in English.

A pool of 67 experts was sent an invitation letter by e-mail with the following information: the SAFEST research team identification, funding of the study, a brief description of the e-Delphi and each of the two rounds, the expected time to be spent during the participation, projected timeline, information on the right to refuse or withdrawfrom the study, associated risks, benefits and compensation, as well as confidentiality information and contacts.

Further iterations post-e-Delphi were considered so as to build upon the Final List of Indicators (FLI) that arose from the consensus method.

2.2. Development of the initial list of indicators (ILI)

An umbrella review works under the assumption that multiple systematic reviews can be amalgamated into a cohesive, unified form (37). In this case, thorough search was conducted across multiple databases: PubMed, EMBASE, Web of Science (Core Collection), Scopus, The Cochrane Library (Cochrane Database of Systematic Reviews), Cumulative Index to Nursing and Allied Health Literature, and COMET Initiative database.

The inclusion criteria for the umbrella review comprised quantitative, qualitative, and mixed-methods systematic reviews that addressed indicators and/or indicator measurement instruments developed within the context of CMS. Systematic reviews that did not provide access to the list of indicators and those that were not performed within the context of CMS were excluded.

A reference list screening was also performed so as to include additional relevant systematic reviews, complementing the electronic search (38). To identify eligible systematic reviews, two reviewers independently screened the titles and abstracts. This was followed by a thorough examination of full-text articles that qualified from the titles and abstracts screening. The process was supported by the utilization of the Rayyan online software. One of the reviewers conducted data extraction using a previously tested and standardized extraction form, while another author reviewed the extracted data. The extracted information encompassed citation details, the type of review, the methodology employed to generate the initial list of indicators, as well as the names, definitions, and measurement instruments of the indicators if they were available. In the event of any discrepancies between the reviewers, consensus was sought, or a third appraiser was consulted to resolve any disagreements.

In order to achieve the consolidated initial list, additional indicators not identified by WP6 were added that derived from a systematic review of Clinical Practice Guidelines conducted by WP2

and from an umbrella review of non-clinical interventions to enhance perioperative patient safety by WP4 (**Figure 1**).

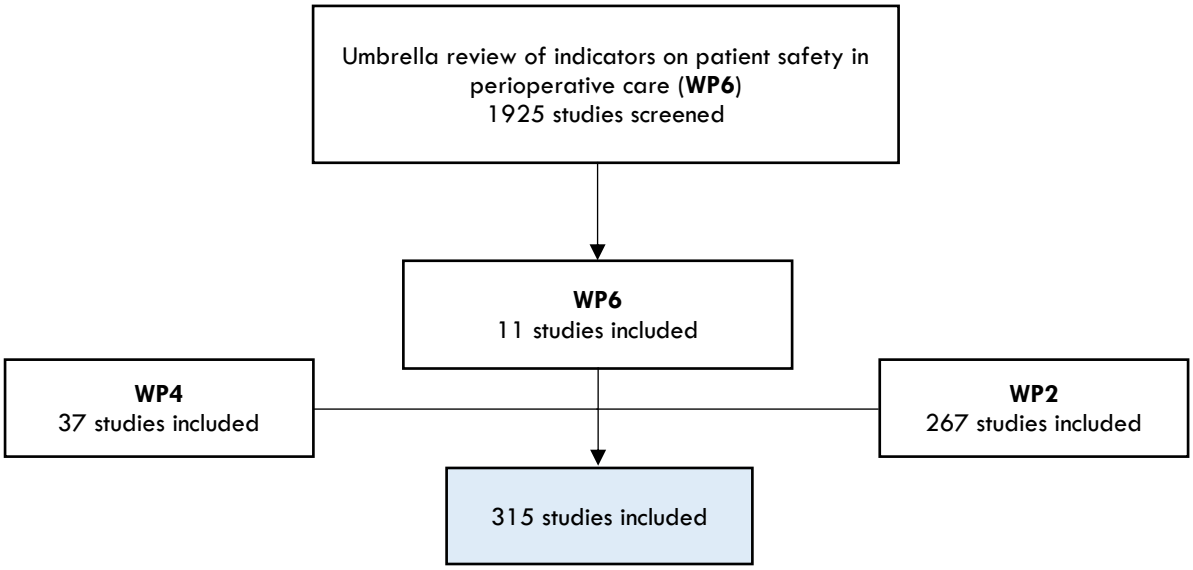

Figure 1. Flow diagram of the systematic reviews conducted towards a consolidated initial list of indicators.

All protocols for the reviews have been registered in PROSPERO. WP6 umbrella review of outcomes and outcome measurement instruments for patient safety in perioperative care has been registered under CRD42022362921. WP2 systematic review has been registered under CRD42022347449 and WP4 umbrella review has been registered under CRD42023397419.

2.3. Achievement of consensus through an e-Delphi method

In the second stage of the process, a consensus panel experts' group was assembled in order to prioritize the initial list of indicators based on the Delphi approach.

Although evidence on how to optimize patient engagement in research is still lacking, the existing body of knowledge is enough to underline the importance of pursuing such participation (39).

The e-Delphi is a modern approach aiming to digitize the Delphi method to enhance its effectiveness in coordinating collective and diverse group thoughts, while leveraging the aforementioned methodological benefits. By utilizing an Internet-based platform, the e-Delphi method streamlines communication between the researcher and the panel of experts, providing organization and control (40).

The e-Delphi used in this study involved two rounds and was conducted as an online survey on the Welphi® platform. The survey was distributed to the experts through personalized links automatically generated by the platform. Participants were able to complete the survey using any internet-connected device, such as mobile phones, laptops, or computers. To assist the consensus panel in this e-Delphi Technique, the WP6 researcher team developed a user manual. This manual provided guidance on the rationale and development of the CMS within the SAFEST project, an overview of the e-Delphi methodology and structure, instructions for using the platform, and the procedures related to the e-Delphi method.

Participants were asked to rank each indicator regarding its perceived importance and feasibility of measurement (from now on, "feasibility"). A Likert 9-point scale was used, where the 1 meant the expert attributed no importance at all to the indicator or considered it to be extremely difficult to measure while value 9 was to be picked whenever the expert considered the indicator critical for inclusion or extremely easy to measure (**Table 1**).

Table 1. Likert scale rating and corresponding descriptive values by importance and feasibility.

Value in Likert-scale	Importance	Feasibility
1	Not important at all	Extremely difficult to measure
2	Not important	Very difficult to measure
3	Not that important	Difficult to measure
4	Slightly important, but not critical	Measurable, but with important difficulties
5	Moderately important, but not critical	Measurable, but with difficulties
6	Important, but not critical	Measurable, but with minor difficulties
7	Very important	Easy to measure
8	Extremely important	Very easy to measure
9	Critical for inclusion	Extremely easy to measure

The expert consensus process was conducted in two rounds between February and May 2023. Round 1 took place from February 24th to March 17th. Round 2 occurred between April 18th and May 2nd. Participants received summaries of key findings and the full list of indicators, including descriptive statistics and anonymized comments at the conclusion of each round. The research team reviewed indicators based on expert feedback and revised names,

definitions, and descriptions between rounds. After Round 1, two new indicators were added, 11 were merged, and four removed based on expert suggestions. Reminders were issued to experts during each round.

Consensus criteria was set for both inclusion and exclusion across round 1 and round 2 of the e-Delphi. If an indicator was to be scored between 7 and 9 by 75% of the experts and between 1 and 3 by 15% or less of experts, it was considered to include. If an indicator was to be scored between 1 and 3 by 75% of the experts and between 7 and 9 by 15% or less of experts, it was considered to exclude.

Indicators achieving consensus for inclusion according to the above criteria and without the need for rewording were included in the final list of indicators (FLI).

2.4. Reaching the final list of indicators

This list of indicators was the result of the e-Delphi. As a first programmed iteration beyond the scope of the e-Delphi, stakeholders were divided in small groups in a CMS Consensus Conference to further discuss the FLI. This conference was held on May 11, 2023 and also served as a platform to discuss barriers and facilitators associated with measuring indicators that are particularly valued by patients and patients' representatives.

3. Results

The indicators initially identified underwent a refinement process that encompassed rewording, merging and deduplication. This was executed by the research team, resulting in a consolidated initial list of 247 indicators.

3.1. e-Delphi Round one findings

The first round of the e-Delphi survey was sent to 67 experts, including 60 patient safety / perioperative professionals across healthcare fields and seven patients or patient representatives. From the initial pool of 67 experts contacted initially, 48 (72%) completed the survey, seven (10%) started but did not finish, and 12 (18%) did not begin. The eight incomplete e-Delphi responses were analyzed, handling missing data through complete-case analysis.

In terms of gender balance, from the group of experts that started the e-Delphi, approximately 55% (n=31) identified themselves as male, while 39,2% (n=22) identified themselves as female. The dominant age group was 55-74 (n=29; 51.8%), while the 34-54 age group was

the second most prevalent (n=29; 39.3%). Most experts had post-graduate education (n=50; 89.3%).

Round 1 respondents represented diverse expert categories. Among them, healthcare professionals (n=43; 76.8%), patients / patient representatives (n=6; 10.7%), government agency representatives (n=3; 5.4%), researchers (n= 2; 3.6%), policymakers (n=1; 1.8%), and private sector reps (n=1; 1.8%). Most healthcare experts were doctors and nurses specialized in anesthesiology and surgery. The patients had surgery themselves or a direct family member did within the last 5 years. Further detailed information on the profile of the experts can be found in S1 table.

From the 247 initial consolidated list of indicators included in Round 1, 40 (16.2%) reached a consensus to include regarding both importance and feasibility.

The distribution of the consensual round 1 indicators by perioperative period and Donabedian's quality of care conceptual model Structure – Process – Outcome subgroups is displayed below in **Table 2**.

Table 2. Distribution of indicators that achieved round 1 consensus by perioperative period and SPO subgroup.

Perioperative period	Preoperative		Intraoperative		Postoperative		Mixed	
S-P-O	n	%	n	%	n	%	n	%
Structure	2	0.8	11	4.5	2	0.8	3	1.2
Process	2	0.8	6	2.4	0	0.0	1	0.4
Outcome	1	0.4	3	1.2	4	1.6	5	2.0

Of these 40 indicators, only nine had no additional comments from experts in Round 1. From these nine, two were reworded based on comments on different indicators. This resulted in eight indicators being immediately endorsed to be integrated into the final list of indicators. No indicators reached a consensus to exclude. As such, no indicators were excluded in this round and all those that did not reach consensus moved on to Round two of the e-Delphi.

The five highest rated indicators in terms of importance on round one were “Equipment to administer oxygen to all patients undergoing procedures under sedation by anaesthetists is available”, “Specialised equipment for the management of difficult airways is available where anaesthesia is given”, “A preoperative up to date medication list is available in the medical records”, and “Postoperative stroke (outcome)”. Regarding feasibility, the five highest rated

indicators were “The operating time is recorded”, “Length of surgery”, “Intraoperative blood transfusion”, “Defibrillators with cardiac pacing mode are available”, and “Equipment for fluid and blood warming and rapid transfusion is available”.

3.2. e-Delphi Round two findings

Based on the responses from Round one, the e-Delphi survey was sent out to 55 experts. One of the 56 round one respondents had only completed to the introductory questions and, as such, was excluded from round two. Out of the 55 experts who received the survey, 85.5% (n=47) responded and completed the survey, while 3.6% (n=2) started but did not finish it, and 10.9% (n=6) did not initiate the second round.

The general profile of the respondents of the second round very strongly overlapped with those from the first round in terms of gender, age group, highest level of education completed and role.

212 indicators were evaluated according to its perceived importance and feasibility in round two. 15.6% (n=33) of the indicators reached consensus to include regarding both importance and feasibility, while none were consensually excluded.

From those 212 indicators assessed in round two, higher scores were observed in importance compared to feasibility within both main classifications of perioperative period and Donabedian model category (Figure 2).

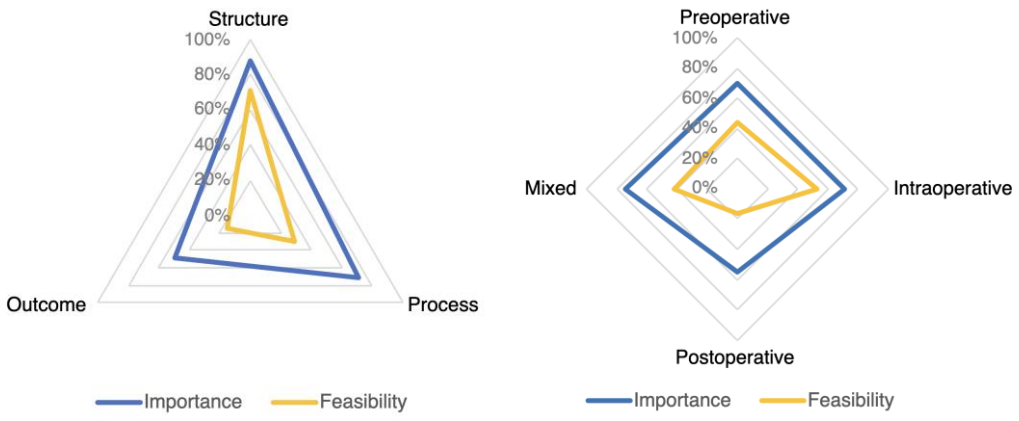

Figure 2. Diagrams with the proportion of indicators that reached consensus according to Donabedian category and perioperative period.

The five highest rated indicators in terms of importance on round two were “Equipment to administer oxygen to all patients undergoing procedures under sedation by anaesthesiologists is available”, “A preoperative up to date medication list is available in the clinical records”,

“There is an internal policy for resuscitation defined and diffused among professionals”, “There is a well-defined internal protocol for major haemorrhage defined, including clinical laboratory and logistic responses, that is diffused among professionals”, “There is a well-defined internal policy that ensures emergency drugs are available where anaesthesia is given and adequately stored defined and this policy is diffused among professionals”. Regarding feasibility, the five highest rated indicators were “There is a well-defined internal protocol for major haemorrhage defined, including clinical laboratory and logistic responses, that is diffused among professionals”, “Fever”, “A preoperative glucose monitoring is conducted in diabetic patients by a knowledgeable and trained professional based on the best available evidence”, “Equipment to administer oxygen to all patients undergoing procedures under sedation by anesthesiologists is available”, and “There is a well-defined internal policy that ensures emergency drugs are available where anaesthesia is given and adequately stored defined and this policy is diffused among professionals”.

Of note, the two highest ranked importance indicators from round two are similar to two of the highest ranked in round one. They are, however, not exactly the same as they have been reworded between rounds.

3.3. Overall e-Delphi findings

Out of the initial list of indicators from the two rounds, 34.1% (n=85) achieved a consensus to be included on the FLI based on the ratings of importance and feasibility. 164 indicators did not reach a consensus to include when considering both importance and feasibility simultaneously. No indicators reached a consensus to exclude.

The number of overall agreed upon indicators to include, categorized by perioperative period and Donabedian's quality of care model Structure-Process-Outcome subgroups, is shown in Figure 3.

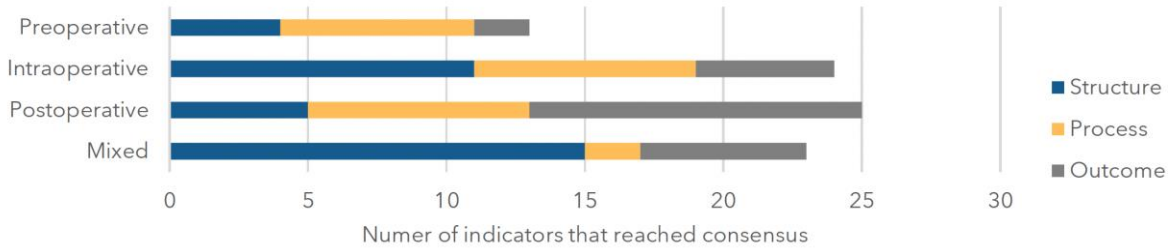

Figure 3. Number of indicators that reached consensus to include after both rounds of the e-Delphi Technique distributed by perioperative period and Donabedian model category.

Different percentage of agreement for consensus was reached across perioperative periods in the ILI. From the initial 32 preoperative indicators, 36 intraoperative indicators, 123 postoperative indicators, 56 mixed period indicators, the following achieved consensus after two rounds of the e-Delphi consensus method: 13 preoperative indicators (~40.6% of the initial number), 24 intraoperative indicators (~66.7%), 25 postoperative indicators (~20.3%) and 23 mixed period indicators (~41.1%). The FLI can be found in S1 appendix.

3.4. Subgroup description: Patients and/or patient representatives and HCPs

By the end of the second round, the top five highest rated indicators by HCPs on average in terms of importance were: “Equipment to administer oxygen to all patients undergoing procedures under sedation by anesthesiologists is available”, “A preoperative up to date medication list is available in the clinical records”, “There is an internal policy for resuscitation defined and diffused among professionals”, “There is a well-defined internal protocol for major haemorrhage defined, including clinical laboratory and logistic responses, that is diffused among professionals“, and “There is a well-defined internal policy that ensures emergency drugs are available where anaesthesia is given and adequately stored defined and this policy is diffused among professionals”. The average rating given by patients or patients’ representatives to each of these indicators is shown in Table 3.

Table 3. Average rating in terms of importance of the top 5 highest HCP-rated indicators as graded by HCP and patients or patient representatives.

Indicator	HCPs average	Patients average
Equipment to administer oxygen to all patients undergoing procedures under sedation by anesthesiologists is available	8.40	8.60
A preoperative up to date medication list is available in the clinical records	8.35	8.60
There is an internal policy for resuscitation defined and diffused among professionals	8.28	8.40
There is a well-defined internal protocol for major haemorrhage defined, including clinical laboratory and logistic responses, that is diffused among professionals	8.24	8.80
There is a well-defined internal policy that ensures emergency drugs are available where anaesthesia is given and adequately stored defined and this policy is diffused among professionals	8.13	8.40

Regarding patient-relevant outcomes, pain-related indicators were assessed in this subgroup description of results. Numeric averages were higher as graded by patients or patients' representatives when compared with the assessment of HCPs regarding the same pain-related indicators.

Table 4. Comparison of the average rating for importance of some patient-experience indicators between patient or patient representatives and HCPs

Indicator	HCPs average	Patients average
Pain associated fear	6.49	8.40
Pain (at rest or on movement)	7.12	8.25
Chronic pain	6.43	8.00
Time to lowest pain score	6.51	7.60
Pain is measured with each set of vital signs	6.76	7.00

4. Discussion

As patient safety gains traction as one of the key pillars of quality of care, governments, national and international organizations in healthcare, and research teams and consortiums like ours have taken steps towards addressing the burden of unsafe, harmful care.

De Vries et al. have clearly reported the relevance of developing comprehensive, multidisciplinary list of surgical safety indicators in the reduction of surgical complications and mortality in hospitals (41). This is in line with the dominant body of evidence regarding the importance of developing checklists and CMS for specific areas of care (11,12). Building upon this idea, our work consisted of the design and operationalization of a two-round e-Delphi consensus method, through which we sought to discern the indicators that garnered agreement among a heterogenous expert panel, encompassing healthcare professionals with diverse backgrounds and expertise. Elaborating on the results of an umbrella review (a systematic review of systematic reviews), the starting point was as broad as possible without compromising the expected compliance of experts in performing the required tasks.

One important limitation of using systematic reviews as the basis of a consensus work, as posed by Bampoe et al. (42), is that this methodology only allows the identification and subsequent consensualization of existing clinical indicators, rather than the development of novel ones. This issue was addressed by allowing experts in our study to propose new indicators during round one of the e-Delphi. As a result, two new indicators ("Patient

satisfaction with perisurgical education" and "Availability of Written or Audiovisual Patient Education Material") have been included by suggestion of the experts. Both served to bridge the gap regarding patient-related indicators found in the literature.

One other area of methodology worth discussing would be the number of Delphi rounds utilized. Regarding this issue, Erffmeyer et al. stated that, although some authors have considered otherwise, the use of two rounds might be insufficient to achieve stability in iterations (43). In our study, after arriving at an ILI, a crucial decision to be made on this process regarding the consensus criteria to be used was pending. As previously alluded to, currently there are still no universally accepted parameters for considering consensus has been achieved in a Delphi study (32,33). Moreover, even though Erffmeyer's perspective on the stability in iterations is a pertinent one, it remains to be proven that the right number of rounds exist. The chosen method to probe for consensus in the present study revolves around the concept of 'controlled feedback'. This means experts were provided with a statistical analysis of the results from the previous round as well as with any comments made beforehand by peers. That data informed the decision-making of the second and last round (44). While a higher number of rounds might have provided an even more robust set of results, the decision of moving ahead with the preset number of two rounds was made after a discussion within the research team that took into consideration practical and operational arguments. Namely, the burden of work requested from the experts across the different work packages and the timelines of the SAFEST project were crucial in the choice.

Patients' perspectives were highly valued from the research team's perspective. Consequently, their representation during the consensus-achieving process was ensured and specific steps were taken during the design of the umbrella review protocol in order to include Patient-Reported Experience Measures (PREMs) and Patient-Reported Outcome Measures (PROMs) in the ILI. We hypothesize that the anonymity of the feedback across rounds allowed for the neutralization of a potential dominance of the opinion of an experienced clinician or researcher against that from a patient or a patient representative.

Taking into consideration the results obtained, it is worth noting that certain crucial domains, such as quality of life, mental health, and satisfaction, were not assigned consensual priority during the process between both patients and professionals. To ensure that the CMS adequately reflects the needs and perspectives of patients, the research team is contemplating the identification of a patient-core set.

Nonetheless, some takeaways from the comparison of answers between patients and HCPs might be drawn. Interestingly, of the top 5 highest rated indicators by HCPs, none had an

average rating below 8.4 by patients or patients' representatives. This is in line with the idea that patients may acquire some expertise in the area they experienced (45,46).

However, there are some differences that our subgroup analysis captured and that highlight the need to involve patients and/or patients' representatives throughout the whole process of reshaping care towards improved quality and safety: from evidence generation to policymaking. Patients tend to value patient-relevant outcomes (47) higher in terms of importance than HCPs, which is corroborated by the results presented in Table 4.

On the other hand, HCPs value some crucial process indicators higher, likely for being particularly aware of their impact on the provision of safe perisurgical care. This difference might be observed from the average importance ratings of indicators such as "There is number of accredited healthcare professionals, including anaesthetists, surgeons and perioperative specialized nurses that is considered adequate according to national recommendations" (Average patient rating: 6.40 / Average HCP rating: 7.72), "supervisory consultants are freely available to all junior anaesthesiologists" (Average patient rating: 6.80 / Average HCP rating: 7.49), or "The average case volume per surgeon is adequate according to national recommendations" (Average patient rating: 6.60 / Average HCP rating: 7.45).

Overall, the results of the present e-Delphi study paint a nuanced picture. While a substantial proportion of indicators found common ground in terms of importance and feasibility, a considerable number failed to achieve consensus in both, reflecting the subtleties inherent to finding indicators able to evaluate perioperative safety.

The overall higher average scores that almost every indicator received in terms of importance when compared to feasibility (Figure 2) stresses the crucial role that measurement plays when developing a set of indicators.

Capturing the complexity of all the perioperative care needed and valued by patients, HCPs, and other key stakeholders is an intricate endeavor. The acknowledgment of this complexity is why indicators were only considered for the FLI if they achieved consensus for both importance and feasibility.

5. Conclusions

This study's findings contribute to the field of patient safety, specifically in the context of perioperative care. The perspective given by the results of this study will inform future concrete iterations of the SAFEST project and ultimately will support healthcare providers in the tracking, measurement, evaluation and regulation of their surgical activities.

In practice, so as to monitor and evaluate the implementation of the SAFEST strategy and recommendations in ten European hospitals, the consortium partners will curate an actionable subset of indicators from the developed Core Measures Set. Guided self-evaluation and implementation materials will be made available to more than 100 hospitals in Europe and beyond.

To better grasp the complexity of perioperative safety indicators, it is important to include a variety of perspectives from different groups in future studies on this topic. Notably, the results of the present study underline how critical the incorporation of patients and patients' representatives in health services research endeavours is. By involving these diverse perspectives, it is possible to achieve results that more comprehensively express actionable insights.

We hope this work can also help to pave the way for future research that should consolidate valid, co-created, patient-centered and action-oriented sets of indicators on the topic of safe perisurgical care.

6. References

1. World Health Organization. Global patient safety action plan 2021–2030: towards eliminating avoidable harm in health care [Internet]. Geneva: World Health Organization; 2021 [cited 2022 Sep 28]. Available from: https://apps.who.int/iris/handle/10665/343477

2. To Err Is Human: Building a Safer Health System [Internet]. Washington, D.C.: National Academies Press; 2000 [cited 2023 Jun 9]. Available from: http://www.nap.edu/catalog/9728

3. United Nations. Resolution adopted by the General Assembly on 25 September 2015. 2015;

4. Sokol DK. “First do no harm” revisited. BMJ. 2013 Oct 25;347(oct25 3):f6426–f6426.

5. De Vries EN, Ramrattan MA, Smorenburg SM, Gouma DJ, Boermeester MA. The incidence and nature of in-hospital adverse events: a systematic review. Qual Saf Health Care. 2008 Jun 1;17(3):216–23.

6. Sousa P, Uva AS, Serranheira F, Uva MS, Nunes C. Patient and hospital characteristics that influence incidence of adverse events in acute public hospitals in Portugal: a retrospective cohort study. Int J Qual Health Care. 2018 Mar 1;30(2):132–7.

7. Bates DW, Levine DM, Salmasian H, Syrowatka A, Shahian DM, Lipsitz S, et al. The Safety of Inpatient Health Care. N Engl J Med. 2023 Jan 12;388(2):142–53.

8. Jha AK, Larizgoitia I, Audera-Lopez C, Prasopa-Plaizier N, Waters H, Bates DW. The global burden of unsafe medical care: analytic modelling of observational studies. BMJ Qual Saf. 2013 Oct;22(10):809–15.

9. The economics of patient safety in primary and ambulatory care: Flying blind [Internet]. 2018 Nov [cited 2023 Feb 6]. (OECD Health Working Papers; vol. 106). Report No.: 106. Available from: https://www.oecd-ilibrary.org/social-issues-migration-health/the-economics-of-patient-safety-in-primary-and-ambulatory-care_baf425ad-en

10. Pearse RM, Moreno RP, Bauer P, Pelosi P, Metnitz P, Spies C, et al. Mortality after surgery in Europe: a 7 day cohort study. The Lancet. 2012 Sep;380(9847):1059–65.

11. Van Klei WA, Hoff RG, Van Aarnhem EEHL, Simmermacher RKJ, Regli LPE, Kappen TH, et al. Effects of the Introduction of the WHO “Surgical Safety Checklist” on In-Hospital Mortality: A Cohort Study. Ann Surg. 2012 Jan;255(1):44–9.

12. Bainbridge D, Martin J, Arango M, Cheng D. Perioperative and anaesthetic-related mortality in developed and developing countries: a systematic review and meta-analysis. The Lancet. 2012 Sep;380(9847):1075–81.

13. The economics of patient safety: From analysis to action [Internet]. 2022 Aug [cited 2023 Jun 9]. (OECD Health Working Papers; vol. 145). Report No.: 145. Available from: https://www.oecd-ilibrary.org/social-issues-migration-health/the-economics-of-patient-safety_761f2da8-en

14. Clarke M. Standardising outcomes for clinical trials and systematic reviews. Trials. 2007 Dec;8(1):39.

15. Williamson PR, Altman DG, Blazeby JM, Clarke M, Devane D, Gargon E, et al. Developing core outcome sets for clinical trials: issues to consider. Trials. 2012 Dec;13(1):132.

16. Burns DJP, Arora J, Okunade O, Beltrame JF, Bernardez-Pereira S, Crespo-Leiro MG, et al. International Consortium for Health Outcomes Measurement (ICHOM): Standardized Patient-Centered Outcomes Measurement Set for Heart Failure Patients. JACC Heart Fail. 2020 Mar;8(3):212–22.

17. Munblit D, Nicholson T, Akrami A, Apfelbacher C, Chen J, De Groote W, et al. A core outcome set for post-COVID-19 condition in adults for use in clinical practice and research: an international Delphi consensus study. Lancet Respir Med. 2022 Jul;10(7):715–24.

18. Karnik NS, Marsden J, McCluskey C, Boley RA, Bradley KA, Campbell CI, et al. The opioid use disorder core outcomes set (OUD–COS) for treatment research: findings from a Delphi consensus study. Addiction. 2022 Sep;117(9):2438–47.

19. Van Tol RR, Melenhorst J, Dirksen CD, Stassen LPS, Breukink SO. Protocol for the development of a Core Outcome Set (COS) for hemorrhoidal disease: an international Delphi study. Int J Colorectal Dis. 2017 Jul;32(7):1091–4.

20. Ravinskaya M, Verbeek JH, Langendam M, Madan I, Verstappen SMM, Kunz R, et al. Which outcomes should always be measured in intervention studies for improving work participation for people with a health problem? An international multistakeholder Delphi

study to develop a core outcome set for Work participation (COS for Work). BMJ Open. 2023 Feb;13(2):e069174.

21. Mahajan V, Linstone HA, Turoff M. The Delphi Method: Techniques and Applications. J Mark Res. 1976 Aug;13(3):317.

22. Hsu CC, Sandford BA. The Delphi Technique: Making Sense of Consensus. [cited 2023 Jan 18]; Available from: https://scholarworks.umass.edu/pare/vol12/iss1/10/

23. The periodic health examination: 2. 1984 update. Canadian Task Force on the Periodic Health Examination. Can Med Assoc J. 1984 May 15;130(10):1278–85.

24. Stegenga J. Down with the Hierarchies. Topoi. 2014 Oct;33(2):313–22.

25. Jorm AF. Using the Delphi expert consensus method in mental health research. Aust N Z J Psychiatry. 2015 Oct;49(10):887–97.

26. Hill J, Ashken T, West S, Macfarlane AJR, El-Boghdadly K, Albrecht E, et al. Core outcome set for peripheral regional anesthesia research: a systematic review and Delphi study. Reg Anesth Pain Med. 2022 Nov;47(11):691–7.

27. Prins JR, Holvast F, Van 'T Hooft J, Bos AF, Ganzevoort JW, Scherjon SA, et al. Development of a core outcome set for immunomodulation in pregnancy (COSIMPREG): a protocol for a systematic review and Delphi study. BMJ Open. 2018 Aug;8(8):e021619.

28. Mease PJ, Arnold LM, Crofford LJ, Williams DA, Russell IJ, Humphrey L, et al. Identifying the clinical domains of fibromyalgia: Contributions from clinician and patient delphi exercises. Arthritis Rheum. 2008 Jul 15;59(7):952–60.

29. Cheung CK, Zebrack B. What do adolescents and young adults want from cancer resources? Insights from a Delphi panel of AYA patients. Support Care Cancer. 2017 Jan;25(1):119–26.

30. Alahlafi A, Burge S. What should undergraduate medical students know about psoriasis? Involving patients in curriculum development: modified Delphi technique. BMJ. 2005 Mar 19;330(7492):633–6.

31. Rowe G, Wright G. Expert Opinions in Forecasting: The Role of the Delphi Technique. In: Armstrong JS, editor. Principles of Forecasting [Internet]. Boston, MA: Springer US; 2001 [cited 2023 Mar 7]. p. 125–44. (Hillier FS, editor. International Series in Operations

Research & Management Science; vol. 30). Available from: http://link.springer.com/10.1007/978-0-306-47630-3_7

32. Jünger S, Payne SA, Brine J, Radbruch L, Brearley SG. Guidance on Conducting and REporting DElphi Studies (CREDES) in palliative care: Recommendations based on a methodological systematic review. Palliat Med. 2017 Sep;31(8):684–706.

33. Trevelyan EG, Robinson PN. Delphi methodology in health research: how to do it? Eur J Integr Med. 2015 Aug;7(4):423–8.

34. Diamond IR, Grant RC, Feldman BM, Pencharz PB, Ling SC, Moore AM, et al. Defining consensus: a systematic review recommends methodologic criteria for reporting of Delphi studies. J Clin Epidemiol. 2014 Apr;67(4):401–9.

35. Shariff NJ. Utilizing the Delphi Survey Approach: A Review. J Nurs Care [Internet]. 2015 [cited 2023 Nov 30];04(03). Available from: https://www.omicsgroup.org/journals/utilizing-the-delphi-survey-approach-a-review--2167-1168-1000246.php?aid=51677

36. SAFEST. Improving quality and patient safety in surgical care through standardisation and harmonisation of perioperative care in Europe - Extended Leaflet [Internet]. Available from: https://safestsurgery.eu/api/wp-content/uploads/2022/11/SAFEST_EXTENDED_LEAFLET.pdf

37. Ioannidis JPA. Integration of evidence from multiple meta-analyses: a primer on umbrella reviews, treatment networks and multiple treatments meta-analyses. Can Med Assoc J. 2009 Oct 13;181(8):488–93.

38. Horsley T, Dingwall O, Sampson M. Checking reference lists to find additional studies for systematic reviews. Cochrane Methodology Review Group, editor. Cochrane Database Syst Rev [Internet]. 2011 Aug 10 [cited 2023 Jul 9];2011(8). Available from: http://doi.wiley.com/10.1002/14651858.MR000026.pub2

39. Domecq JP, Prutsky G, Elraiyah T, Wang Z, Nabhan M, Shippee N, et al. Patient engagement in research: a systematic review. BMC Health Serv Res. 2014 Feb 26;14(1):89.

40. Donohoe H, Stellefson M, Tennant B. Advantages and Limitations of the e-Delphi Technique: Implications for Health Education Researchers. Am J Health Educ. 2012 Jan;43(1):38–46.

41. De Vries EN, Prins HA, Crolla RMPH, Den Outer AJ, Van Andel G, Van Helden SH, et al. Effect of a Comprehensive Surgical Safety System on Patient Outcomes. N Engl J Med. 2010 Nov 11;363(20):1928–37.

42. Bampoe S, Cook T, Fleisher L, Grocott MPW, Neuman M, Story D, et al. Clinical indicators for reporting the effectiveness of patient quality and safety-related interventions: a protocol of a systematic review and Delphi consensus process as part of the international Standardised Endpoints for Perioperative Medicine initiative (StEP). BMJ Open. 2018 Nov;8(11):e023427.

43. Erffmeyer RC, Erffmeyer ES, Lane IM. The Delphi Technique: An Empirical Evaluation of the Optimal Number of Rounds. Group Organ Stud. 1986 Mar;11(1–2):120–8.

44. Boulkedid R, Abdoul H, Loustau M, Sibony O, Alberti C. Using and reporting the Delphi method for selecting healthcare quality indicators: a systematic review. PloS One. 2011;6(6):e20476.

45. Wilson J. Acknowledging the expertise of patients and their organisations. BMJ. 1999 Sep 18;319(7212):771–4.

46. Norburn L, Thomas L. Expertise, experience, and excellence. Twenty years of patient involvement in health technology assessment at NICE: an evolving story. Int J Technol Assess Health Care. 2021;37(1):e15.

47. Kersting C, Kneer M, Barzel A. Patient-relevant outcomes: what are we talking about? A scoping review to improve conceptual clarity. BMC Health Serv Res. 2020 Dec;20(1):596.

S1 Appendix. Final List of Indicators.

Preoperative Indicators

Structure Indicators
A consultant anaesthesiologist is responsible for leading the anaesthetic preoperative assessment service.
There is an internal policy for preoperative preparation defined and diffused among professionals, including all the following: fasting, investigations, blood typing, thromboprophylaxis, perioperative diabetes management, and allergies.
There is an agreed internal policy for referral pathways to other specialties to expedite further investigations that is diffused among professionals.
There is a defined internal policy to ensure that abnormal results of investigations are flagged to the relevant person in a timely manner defined that is diffused among professionals.
Process Indicators
Preassessment according to local recommendations is conducted at least the day before surgery.

A preoperative up to date medication list is available in the clinical records.
A preoperative glucose monitoring is conducted in diabetic patients by a knowledgeable and trained professional based on the best available evidence.
Re-assess of venous thromboembolism and bleeding risk using risk assessment criteria is conducted on admission and within 24 hours of admission.
Patients assessed to be at risk of venous thromboembolism are offered prophylaxis in accordance with best practices.
The stoma site is marked if indicated.
Risk assessment for pressure ulcers using a standardized scale upon admission is conducted.
Outcome Indicators
Proportion of prospective surgical patients that undergo electrocardiographic (ECG) assessment preoperatively.
Proportion of prospective surgical patients that have their full blood count, coagulation profile and renal function checked preoperatively.

Intraoperative Indicators

Structure Indicators
Properly designed transfer trolleys meet the requirements of the following list: oxygen cylinders, masks, tubing, infusion poles, equipment to secure and support airway and assist ventilation, provision of clamps for drainage tubes, protective sides, head down tilt possible are available.
Equipment to administer oxygen to all patients undergoing procedures under sedation by anaesthetists is available.
Specialised equipment for the management of difficult airways is available where anaesthesia is given.
There is a well-defined internal policy for sedation that includes the training required by the sedation provider, all subspecialty areas and facility specifications and this policy is diffused among professionals.
There is a well-defined internal policy that ensures emergency drugs are available where anaesthesia is given and adequately stored defined and this policy is diffused among professionals.

There is a well-defined internal policy for the management of complications of anesthesiologist procedures and this policy is diffused among professionals.
Devices for maintaining or raising the patient's temperature are available, including control of theatre temperature. [a]
Defibrillators with cardiac pacing mode are available.
Equipment for fluid and blood warming and rapid transfusion is available.
Blood storage facilities are in close proximity to emergency theatres and contain 0 rhesus negative blood.
There is a well-defined internal protocol for major haemorrhage defined, including clinical laboratory and logistic responses, that is diffused among professionals.
Process Indicators
An appropriate antibiotic is given as per local guidelines.
Equipment used to provide anaesthesia, including monitoring equipment, complies with existing local recommendations.
Surgical procedures with predicted mortality >10% are conducted under the direct supervision of a consultant surgeon and anaesthesiologist.
The *WHO Surgical Safety Checklist* checklist is applied.
Intraoperative blood loss is measured and recorded. [a]
Surgical pathology specimens are labelled according to recommendations, including: labelled, filled containers, correct laterality, correct tissue type, patient name, and correct patient name.
The turnover time between cases is measured.
The operating time is recorded. [a]
Outcome Indicators
Failed attempt of endotracheal intubation.
Wrong site surgery. [a]
Intraoperative blood transfusion.
Unanticipated transfusion of any blood products.
Length of surgery.

Postoperative Indicators

Structure Indicators
The Post-Anesthesia Care Unit equipment includes: - At bedside: pulse oximetry, ECG and Noninvasive Blood Pressure Monitoring - Immediately available: capnograph, 12 lead ECG, nerve stimulator, thermometer.

There is an internal system for ordering, storing, recording and auditing controlled drugs (e.g. morphine, fentanyl) in all postoperative areas where they are used.
There is an internal procedure defined for removing endotracheal tubes and supraglottic airways and that is diffused among professionals.
There is an internal policy for a member of the anaesthetic/clinical team to visit patients within 24 hours following the surgery (ASA grade 3,4,5: epidural on ward, invasive monitoring in-situ or as requested by health care worker) defined.
Internal criteria for discharge from recovery ward are defined and diffused among professionals.
Process Indicators
The postoperative morphine consumption at 6, 24 and 48 hours is recorded.
The analgesic supplementation by any route at 24 hours is recorded.
Post-anaesthesia medical records are compliant with local recommendations. The following are recorded: Information about patient evaluation on admission and discharge from Post-Anesthesia Care Unit or admission to the Intensive Care Unit, a time-based record of vital signs and level of consciousness, time-based record of drugs administered, dosage and route of administration, type and counts of intravenous fluids administered, including blood and blood products, and post-anaesthesia visits.
The Post-Anesthesia Care Unit's length of stay is recorded.
Early warning systems are used at ward care.
The discharge destination is recorded.
Medical records are compliant with local recommendations. Information about discharge needs assessment and venous thromboembolism prophylaxis is recorded.
Recovery area complies with local standards.
Outcome Indicators
Mortality.
Fever.
Postoperative sepsis.
Septic Shock.
Postoperative pneumonia.
Antibiotics use.
Unplanned return to operating theatre within hospital stay or intervention in the same area or for a reason related to the original intervention.
Readmission to the Intensive Care Unit or Intermediate Medical Care Unit. [a]
Unplanned readmission to hospital.
Length of stay.

Length of stay in Post-Anaesthesia Recovery Area.
Length of stay in the Intensive Care Unit. [a]

Mixed Perioperative Indicators

Structure Indicators
Alternative language leaflets or videos and interpreters appropriate to the needs of the local population are available to patients and caregivers.
There is an internal policy for senior clinicians to discuss the defined limits of care and resuscitation that is diffused among professionals.
There is a defined internal policy for planned maintenance and replacement programme for anaesthetic equipment defined, including naming a consultant to oversee the provision of anaesthetic equipment that is diffused among professionals.
There is an internal policy for anaesthetic emergencies defined and diffused among professionals.
There is an internal policy for managing morbidly obese patients defined and diffused among professionals.
There is an internal policy for remote site anaesthesia defined and diffused among professionals.
There is an internal policy for critical care referral defined and diffused among professionals.
There is an internal policy for resuscitation defined and diffused among professionals.
There is an internal policy for end of life care defined and diffused among professionals.
There is an internal policy for staff training for both technical and non-technical skills defined, including in resuscitation.
There is an internal policy for the handover of care of the patient from one team to the other throughout the perioperative pathway defined. [a]
There is an internal policy for the management and reporting of adverse events and near miss events relating to the perisurgical period.
Facilities for rest for on-call/on-duty staff are available.
The number of existing theatres (excluding radiology suites, dedicated obstetric, minor operations but including day theatre) is considered adequate according to national recommendations.
There is an internal policy for receiving feedback from patients and caregivers, including complaints, in place.
Process Indicators

Whether pain is measured within the postoperative period, including pre-discharge and after discharge.
The length of stay is measured.
Outcome Indicators
Absence of falls following surgery.
Perioperative hypothermia.
Perioperative hypoglycemic events.
Transfused patients.
Venous thromboembolism prophylaxis.
Anticoagulation therapy. [a]

[a] Indicator included in the FLI after the first round of the eDelphi

Table S1. Detailed information on the characteristics of the experts.

	n	%
Gender		
Female	22	39,3%
Male	31	55,4%
Non-binary	0	0,0%
Prefer not to answer	3	5,4%
Total	56	100,0%
Age group		
18-34	2	3,6%
35-54	22	39,3%
55-74	29	51,8%
75+	1	1,8%
Prefer not to answer	2	3,6%

Total	56	100,0%
Highest Level of Education completed		
Postgraduate education	50	89,3%
Tertiary (higher) education	4	7,1%
Secondary school	2	3,6%
Primary school	0	0,0%
None / Incomplete primary school	0	0,0%
Prefer not to answer	0	0,0%
Total	56	100,0%
Role		
Healthcare professional	43	76,8%
Profession		
Medical Doctor	33	76,7%
Nurse	5	11,6%
Quality expert	2	4,7%
Hospital manager	1	2,3%
Physical therapist	1	2,3%
Other	1	2,3%
Prefer not to answer	0	0,0%
Total	43	100,0%
Area of expertise		
Anaesthesiology	20	46,5%
Surgery	9	20,9%
Primary Care	3	7,0%
Public Health	3	7,0%
Haematology	2	4,7%
Rehabilitation	1	2,3%
Other	5	11,6%
Geriatrics	0	0,0%
Prefer not to answer	0	0,0%
Total	43	100,0%
Patients and/or patient representative	6	10,7%
Submitted to surgery ≤ 5 years ago		

Yes	2	33,3%
No	1	16,7%
No, but a direct family member was	3	50,0%
Prefer not to answer	0	0,0%
Total	6	100,0%
Governmental agency representative	3	5,4%
Methodologist/Researcher	2	3,6%
Policymaker	1	1,8%
Private sector representative	1	1,8%
Regulatory agencies representative	0	0,0%
Guideline developer	0	0,0%
Other	0	0,0%
Prefer not to answer	0	0,0%
Total	56	100,0%
Any Conflict of Interest declared		
Yes	4	7,1%
No	52	92,9%
Total	56	100,0%