Supplementary Table 2 IDEAL Checklists for IDEAL Stages 1, 2a, 2b, and 4—examples Included to Demonstrate Good Reporting for Key/Unique Items in Each IDEAL Stage

|  |  |  |
| --- | --- | --- |
|  | Item | Checklist Item and example for key IDEAL items |
| **STAGE 1** |  |  |
| Title and Abstract | 1a | **Identify the technique or device in the title, including IDEAL Stage 1 or ‘first in human’ in the title or abstract.** |
|  | 1b | **Provide a structured summary of background, methods, results, and conclusions.** |
| **Introduction** |  |  |
| Background and objectives | 2a | **Review of existing scientific literature, providing a clear explanation of the rationale for the new technique, including unmet clinical need.** |
| 2b | **Details of pre-clinical development of the technique, including assessment of risks of failure and analysis of efforts to avoid harm.\***Example**:** “Central to achieving successful graft function is the ability to minimize warm ischemia]. Thus a critical part of the phase 0 study was the development of techniques to cool the graft during RKT, particularly during the potentially longer initial cases.”1  |
| **Methods** |  |  |
| Design | 3 | **Description of study design (e.g. case report or very small case series).** Example: “Phase 0 (simulation) studies included the establishment of techniques for pelvic cooling, graft placement in a robotic prostatectomy model, and simulation of the RKT procedure in a cadaveric model. Phase 1 (innovation) studies began in January 2013 and involved treatment of a highly selective small group of patients (n = 7), using the principles utilized in the phase 0 studies, at a tertiary referral center.” 1 |
| Participants | 4a | **Transparent account of patient selection, with explicit detail about inclusion and exclusion criteria.** Example: “We assessed 16 patients, of whom 12 were placed on the waiting list; of the four patients excluded, two had residual phalluses that were deemed adequate for all penile functions, making risk–benefit ratios unfavourable, and two had psychiatric diseases that precluded safe transplantation….Suitable participants underwent tissue typing, blood grouping, and screens for hepatitis B virus, cytomegalovirus, and HIV” “Patients with HIV or hepatitis B virus infections who failed to develop immunity to the hepatitis B virus were excluded. Other exclusion criteria were any active bacterial infection…”2 |
| 4b | **Informed consent process described, including explanation of risks and acknowledgement of level of experience with technique/device. If informed consent is not obtained due to unplanned technique or modification, describe the discussion with the patient after the innovation. occurred.** Example: “All patients gave their written informed consent to test our navigation system for renal colleting system puncture after the risks, benefits, and alternatives were discussed. Institutional review board approval was obtained before the start of the study. Patients were specifically informed that this was first clinical application of this novel system.”3 |
|  | 4c | **Setting, location, and timeframe of when and where the novel technique was performed, including hospital characteristics and appropriate details regarding the operator/team (e.g. experience).** |
| Intervention | 5a | **Clear and detailed description of the new technique/device, including necessary pre- and post-procedure care.** Example: “The accompanying video illustrates the technology and provides a step-by-step description of the procedure. Under general anesthesia, the patient is placed in the supine position to allow a combined approach with flexible ureterorenoscopy and percutaneous.”3 |
|  | 5b | **Patient safety monitoring methods and safeguards.** |
| Outcomes | 6 | **Description of outcome measure(s) selected and how they were assessed, including patient reported outcome measures, if appropriate, utilising those measures that are standardised and validated, when available and applicable. When these are not available, provide rationale for the outcome measure(s) used.** Example: “Table 1, Table 2 summarize the demographic, perioperative, and postoperative data of the initial seven patients.”“Mean intraoperative kidney surface temperature measured just before reperfusion was 22.5oC and inversely related to the rewarming time (Fig. 5b)….. No patient developed systemic hypothermia. Mean serum creatinine on postoperative day (POD) 1 and 7 was 2.2 mg/dl and 1.2 mg/dl, respectively (Fig. 5c). No patient required dialysis.” 3 |
| **Results** |  |  |
| Baseline Data | 7 | **Baseline demographic and clinical characteristics for each patient. Include how many patients were assessed for treatment and a description of which patients were included, excluded, or refused, and why (to be displayed in a flow diagram format, when appropriate).** Example: “We screened all potential participants physically and psychologically. We assessed 16 patients, of whom 12 were placed on the waiting list; of the four patients excluded, two had residual phalluses that were deemed adequate for all penile functions, making risk–benefit ratios unfavourable, and two had psychiatric diseases that precluded safe transplantation. “Patients with HIV or hepatitis B virus infections who failed to develop immunity to the hepatitis B virus were excluded. Other exclusion criteria were any active bacterial infection (including tuberculosis), any evidence of psychiatric disease or history of substance abuse….” 2 |
| Intervention | 8 | **Technical feasibility of technique, including visual aids (e.g. photographs, videos, etc) when available.** Example: “2.2. Demonstration of ability to achieve effective pelvic cooling: Central to achieving successful graft function is the ability to minimize warm ischemia. Thus a critical part of the phase 0 study was the development of techniques to cool the graft during RKT, particularly during the potentially longer initial cases. This was established at the Vattikuti Urology Institute during robot-assisted (laparoscopic) radical prostatectomy (RARP) in >300 patients, where 240–300 ml of ice slush was introduced through the GelPOINT.” 1 |
| Outcomes | 9 | **Appropriate clinical outcomes, including patient-reported outcome measures, when applicable.** Example: “Clinical Outcomes: At the time of study completion, the median clinical follow-up time was 13 months (IQR 10–17 months)…. The mean BPI severity score at baseline was 3.8 (range 1–7) and decreased to 2.7 (range 0–5) at 4 weeks post-treatment, with all but one patient reporting a decrease in pain. All patients experienced a partial pain response during follow-up according to the consensus criteria. At 4 weeks post-treatment substantial improvements were reported in all domains of the SOSGOQ, BM-22, QLQ-C15, and the EQ-5D with further improvement over time.”4 |
| Harms | 10 | **Transparent account of all harms or unintended effects reported for each patient.** Example: “Adverse Events: In the interval between SBRT and surgery we observed two grade 1 adverse events; nausea and radiation dermatitis (erythema). Cement leakage outside the vertebral body was observed intra-operatively with fluoroscopy in three patients. In one patient, this caused grade 3 radiculitis due to compression of the exiting left L3 nerve root requiring re-operation to decompress the root…”4 |
| **Discussion** |  |  |
| Stage End-Points | 11 | **Author’s overall appraisal of the new technique, including discussion of risks and harms reported and suggestions to avoid them in future cases based on initial experience.** Example: “Our findings show that penile allotransplantation is feasible and can result in restoration of sexual function, penile sensation, and normal urination.” “Although the results of the transplant have been rewarding, we are aware that the potential long-term consequences of sustained immunosuppression remain a risk and will be monitored for carefully.” 2 |
| Conclusions | 12 | **Conclusions and relevance, including plans to progress to future IDEAL stages, or plans to discontinue further research.** Example: “In conclusion, this study demonstrated the safety and feasibility of SBRT, with active sparing of the surgical area, followed by surgical stabilization within 24 h for the treatment of symptomatic unstable spinal metastases with none of the patients demonstrating disturbed wound healing.“ ” An IDEAL stage IIb study is currently planned to evaluate the effectiveness of the new treatment strategy and to obtain additional data to potentially change the standard of care for patients with symptomatic unstable spinal metastases.” 4 |
| **Other Information** |  |  |
| Protocol | 13 | **Please quote reference or DOI if a protocol was written in advance and made available. If a protocol was not made available, consider including as a supplement if the journal allows** |
| Ethics | 14 | **Reference to ethical approvals obtained, and independent oversight, if applicable** |
| Funding | 15 | **Sources of funding and support, role of funders, and other conflicts of interest** |
| Regulatory Approvals | 16 | **Regulatory approvals being sought or obtained (e.g. CE Marking, FDA approval, etc) including the date of approval, if applicable** |
|  |  |  |
| **STAGE 2a** |  |  |
| Title and Abstract | 1a | **Identification as a prospective case series of a novel technique in the title, including the IDEAL stage in the title or abstract.** |
|  | 1b | **Provide a structured summary of background, methods, results, and conclusions.** |
| **Introduction** |  |  |
| Background and Objectives | 2a | **Review of existing scientific literature, including reference to IDEAL Stage 1 and 2a reports in previous publications, if applicable.** |
|  | 2b | **Specific objectives stated, including refining the technique and progressing toward stability.** Example: “The aim of this study was to demonstrate feasibility; assess the safety and the short-term effects of the surgical isolation of the arterial blood supply of the ipsilateral liver lobe containinga large HCC…. surgically, by assessing the operative mortality, complications and tumor response at the initial postoperativethirty days' following this procedure**.“ ”**This innovative procedure is reported as development stage 2a till it becomes stable and a real comparative design will subsequently be instituted.”5 |
| **Methods** |  |  |
| Design | 3 | **Description of study design (e.g. sequentially reported prospective case series).** Example: “This is a prospective; single arm; single center and stage 2a development, IDEAL (Idea, Development, Exploration, Assessment, and Long-term monitoring) case series.”5 |
| Participants | 4a | **Detailed account of patient inclusion and exclusion criteria.** Example: See Table 1: Inclusion and exclusion criteria for the study. Inclusion criteria include age > 18 years, Body mass index <35 kg/m2, Suitable for minimally invasive surgery, Willingness to participate as demonstrated by giving informed consent, Adequate exposure to the surgical site according to the surgeon’s judgment to proceed with surgery….”6 |
|  | 4b | **Informed consent process described, including explanation of risks and acknowledgement of level of experience with technique/device.** |
|  | 4c | **Setting, location, and timeframe of recruitment and follow-up, including when and where the data were collected, as well as hospital characteristics and appropriate details regarding the operator/team (e.g. prior experience with novel technique).** |
| Intervention | 5a | **Clear and detailed description of (or reference to) planned technique, including necessary pre-operative and post-operative care.** Example: Donor surgery involved a midline incision from the pubic bone to the umbilicus, followed by isolation of the uterus with long vascular pedicles consisting of the bilateral uterine arteries and veins up to and including parts of the internal iliac vessels (Fig. 1A). Substantial parts of the round ligaments and the sacrouterine ligaments, as well as an extensive sheet of the bladder peritoneum, were preserved on the graft side to enable stable fixation of the uterus in the recipient. Bilateral salpingectomy was performed, preserving the uterine branch of the utero-ovarian vein.”7 |
|  | 5b | **Patient safety monitoring methods and safeguards.** |
| Outcomes | 6 | **Description of outcome measure(s) selected and how they will be assessed, including patient reported outcome measures, when appropriate, utilising those measures that are standardised and validated, when available and applicable. When these are not available, provide rationale for the outcome measure(s) used.** Example:Perioperative complications including intraoperative complications and all complications occurring during the hospital stay or within 30 days after discharge were graded according to the Clavien-Dindo classification. Complications of Clavien-Dindo grade III (those requiring surgical, endoscopic, or radiologic intervention) or above were regarded as major complications.” 6 |
| **Results** |  |  |
| Baseline data | 7 | **Patient baseline demographic and clinical characteristics, including how many patients were assessed for treatment and a description of which patients were included, excluded, or refused, and why (to be displayed in a flow diagram format, when appropriate).** Example: “Seven donor/recipient pairs were ABO identical, and two were compatible. Both cytotoxic and flow-cytometric crossmatch tests were negative in all patients, and no patients had HLA antibodies. The degree of HLA mismatch among the donor and recipient pairs varied from 1/0 to 3/2. Table 3: Surgical, anesthesiologic, and hospitalization parameters. The characteristics of recipients and donors are presented in Table 2…..”7  |
| Intervention | 8 | **Transparent reporting of all cases in the sequence they were performed, clearly indicating when and why modifications to the technique took place, including visual aids of the technique and modifications (e.g. photographs, videos, etc) when available.** Example:“After case 3, we changed from using robotic to conventional laparoscopy because of the disadvantages associated with using the robot in this phase: more difficulty and danger in visceral manipulation over wide angles and frequent need for staplers, clipping and sealing devices which are managed by the assistant. Four of the next 6 cases had mid-thoracic lesions and were scheduled for initial robotic thoracoscopy, conventional laparoscopy and neck anastomosis. However in case 5 the robot was not available ….”8 (For excellent visual aides of a surgical technique, see Menon).1  |
| Outcomes | 9 | **Technical, clinical and patient-reported outcomes described for each patient, with all available outcome data incorporated into a comprehensive table or graph, whenever possible, to allow for the relationship to be clearly visualized between technique modifications and outcomes.** Example: See the visual aids in paper illustrating whether there were temporal relationships between technique modifications and patient outcomes. When specific technical changes were implemented are denoted on the x axis while complications are denoted above by downward pointing arrows.” 8 |
| Harms | 10 | **Transparent account of all harms or unintended effects reported for each patient.** Example: See item 12 in stage 2B. |
| **Discussion** |  |  |
| Interpretation | 11 | **Analysis of technique development, including consistency of results and a balanced discussion of benefits and harms.** Example: **“**this safety and feasibility trial of the da Vinci SP demonstrates that the device is likely safe and that it is feasible in performing transoral surgery to access the nasopharynx, oropharynx, larynx and hypopharyngeal region for the treatment of both benign and malignant lesions. Importantly, this has been achieved in a predominantly Southern Chinese population utilizing a Crowe-Davis oral retractor to overcome the particular, smaller cranio-facial measurements in our population as compared to Caucasians….There are some limitations with the current system, despite the benefit with the extra arm there is less working space for an assistant to, in particular, aid with suction. However, this can be largely overcome with a suction catheter placed through the nasal cavity into the desired location. Another limitation is the lack of bone instrumentation to resect, in particular tumours of the nasopharynx and skull base. Finally, tailor made instrumentation for finer dissections are needed to expand the performance for the system to be used on the glottic larynx.” 6 |
| Limitations | 12 | **Study limitations, addressing sources of potential bias.** Example:See item 14 in Stage 2B |
| *Stage End-Points* | 13a | **Have the technique and outcomes reached stability in the hands of the current team (e.g. there is no intent to make further major modifications to the technique, and patient outcomes are stable)? Include an explanation of how you determined stability.** Example: Blazeby, et al, report the instability of the technique and resulting decision: “…stage 2a (modiﬁcation of the technique to two-phase MIO) was abandoned following consecutive technical complications.”9  |
|  | 13b | **Discussion of whether the technique is ready for evaluation in a prospective, multi-centre IDEAL Stage 2b study, and identification of indications for the technique.** Example:Our phase 2a studies demonstrate that RKT with regional hypothermia is a safe and reproducible procedure in the hands of a surgical team with experience in robotic surgery and conventional KT. The next step in the development of this procedure is a comparative analysis of outcomes following open KT and RKT in matched patients.”1 |
| Conclusions | 14 | **Conclusions and relevance, including plans to progress to future IDEAL stages, if applicable.** Example: The primary results of this study showed that hepatic artery ligation with extrahepatic collaterals division of the liver lobe diseased with large HCC as a new alternative treatment option is feasible, safe and produces good tumor response as a short-term evidence base. A similar-design long term study will be needed to asssess the stability of the innovation on the long term. Evaluation of other events such as survival (which is the endpoint treatment of HCC) and localrecurrence and disease free survival will no doubt, be recommended.5 |
| **Other Information** |  |  |
| Protocol | 15e | **Please quote reference or DOI if a protocol was written in advance and made available. If a protocol was not made available, consider including as a supplement if the journal allows.** |
| Ethics | 15f | **Reference to ethical approvals obtained, and independent oversight, if applicable.** |
| Funding | 15g | **Sources of funding and support, role of funders, and other conflicts of interest.**  |
| Regulatory Approvals | 16 | **Regulatory approvals being sought or obtained (e.g. CE Marking, FDA approval, etc) including the date of approval, if applicable.**  |
|  |  |  |
| **STAGE 2b** |  |  |
| Title and Abstract | 1a | **Identify the novel technique/device being investigated and the type of study conducted (e.g. multi-centre, prospective cohort or feasibility RCT), including the IDEAL stage in the title or abstract.** |
|  | 1b | **Provide a structured summary of background, methods, results, and conclusions.** |
| **Introduction** |  |  |
| Background and Objectives | 2a | **Review of existing scientific literature, including reference to IDEAL Stage 1 and 2a reports in previous publications, if applicable.** |
|  | 2b | **Specific objectives, including reaching consensus on issues requiring resolution in order to determine whether an RCT is appropriate or feasible, and to define the RCT question if it is.** Example: “The aim of Pre-BRA is to determine the feasibility of using mixed-methods within an IDEAL 2a/2b study to explore the short-term safety and effectiveness of PPBR and determine when it is sufficiently stable for formal evaluation in a definitive RCT. Specific objectives:To establish the short-term safety (implant loss; infection; reoperation and readmission at 3 months) and effectiveness (patient-reported outcomes using BREAST-Q at 3 and 18 months) of PPBR compared with published national standards from the National Mastectomy and Breast Reconstruction Audit31 and iBRA7 studies.”10 |
| **Methods** |  |  |
| Design | 3 | **Description of multi-centre study design, with prospective collection of standard data across centres.** Example” **“**We conducted a 20‐centre non‐randomised controlled prospective cohort study of treatments for uterine fibroids, in which patients were fully informed about the treatment options and then chose whether to receive hysterectomy, myomectomy or HIFU therapy. Setting and data collection:The centres of clinical investigation were selected from respondents to a nationwide call for participation in the trial. Twenty centres were selected based on having the required ultrasound therapeutic devices in place and conforming to a uniform treatment protocol. Recruitment opened in March 2011 and closed in December 2013.”11 |
| Participants | 4a | **Detailed account of patient inclusion and exclusion criteria, defining any specific recognized patient sub-groups or technique/device variants included for which there is not full consensus over their eligibility.** Example: “Inclusion criteria:Consecutive women aged 16 or over who require a mastectomy for breast cancer or risk-reduction who elect to undergo IBBR are considered technically suitable for PPBR by their surgeon and consent to undergo the procedure are eligible to participate in the study” “Exclusion criteria:Absolute and relative exclusion criteria will be used to reflect current practice and account for the variable experience of surgeons performing PPBR across the UK. Absolute exclusion criteria will be applied to all patients. Absolute exclusion criteria: Patients assessed as having insufficient soft tissue coverage for PPBR…”10 |
|  | 4b | **Informed consent process described, including explanation of risks and acknowledgement of level of experience with technique/device.** |
|  | 4c | **Setting, location, and timeframe of recruitment and follow-up, including when and where the data were collected, as well as hospital characteristics and appropriate details regarding the operator/team (e.g. prior experience with novel technique).** |
| Intervention | 5a | **Clear and detailed description of technique, or reference to it, including an assessment measure for quality of performance for operators/teams.** Example: “Irreversible Electroporation Ablation (Study Intervention):In this study, IRE ablation is performed using the NanoKnife IRE device (AngioDynamics Inc) ([Figure 2](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5334515/figure/figure2/), A), also registered as the HVP-01 electroporation system. The IRE system contains a low energy direct current generator, a foot switch, and 19G monopolar needle electrodes (15 or 25 cm length).”” The IRE procedure will be performed at the radiology department CT room with the patient under general anesthesia including deep muscle relaxation to prevent severe muscle contraction… guided by CT and accompanied by an external spacer for fixation, needle electrodes will be placed ([Figure 2](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5334515/figure/figure2/), B).”12 |
|  | 5b | **Description or reference to statistical learning curve assessment of operators/team using pre-defined objective quality metrics.**“Evaluation of learning curves: Procedure time and percent NPV on enhanced MRI at 4 weeks were measured, to assess the quality of treatment delivery and assess learning curves. For each centre, %NPV was analysed sequentially using the LC‐CUSUM method to identify the point at which a stable satisfactory performance could be demonstrated.15,16 The target for %NPV was set at 70%, and acceptable and unacceptable failure rates were set at 0.1 and 0.2, i.e. 10% and 20% failure to achieve the 70% target.”11 |
|  | 5c | **Patient safety monitoring methods and safeguards.** |
| Outcomes | 6 | **Description of pre-specified primary and secondary outcome measures selected and how they will be assessed, including patient reported outcome measures, when appropriate, utilising standardised and validated core outcomes sets, when available and applicable. When these are not available, provide rationale for the outcome measure(s) used.** Example:“In the present context, a prospective database was set up to record patient details including age, sex, body mass index (BMI), preoperative chemotherapy regimen, cancer stage, site and type. Process measures included duration of surgery (defined as knife to skin to completion of wound closure), intraoperative blood loss (recorded by the surgeon from in‐theatre assessment of blood loss) and pathological disease stage.7 Short‐term measures included length of hospital stay (defined as the number of days spent in hospital from day of operation to discharge), in‐hospital morbidity and in‐hospital mortality. Morbidity was graded according to the Clavien–Dindo classification, which grades complications based on the intervention required for their treatment…”9 |
| Statistical Methods | 7 | **Statistical methods used to describe baseline characteristics and evaluate primary and secondary outcomes, when appropriate, including methods for additional analysis (e.g. learning curve analysis, pre-specified subgroup analysis, evaluation of known confounders).** Example: “Group comparisons were carried out using the chi‐square test for comparison of proportions, two‐sided Student's *t*‐test for continuous parametric variables and the Kruskal–Wallis (3 group) and Wilcoxon (2 group) tests for categorical non‐parametric data comparison. To adjust for the effects of inequalities between treatment groups, multiple regression modelling was performed to identify the influence of preselected covariates on the dependent variable. This process was carried out for each of the dimensions of SF‐36….”11 |
| Stakeholder Values | 8 | **Describe or reference attempts to evaluate patient and surgeon preferences and values relevant to future RCT trial design and conduct, including any qualitative work done to ascertain views about randomization.** Example:“Surgeon engagement will be used to assess the feasibility of using mixed-methods to explore learning and intervention stability in the study. At study entry, the number of surgeons completing the surgeon questionnaire; the number of surgeons consenting to be contacted for interview and the number of surgeons interviewed will be compared. For modifications/learning arising from complications, the numbers of events (modifications and complications) reported will be compared with numbers of free-text responses entered; the number of surgeons consenting to interview and the actual number interviewed. The content of the free-text responses and interviews will also be assessed…”10 |
| **Results** |  |  |
| Baseline data | 9 | **Patient baseline demographic and clinical characteristics, including how many patients were assessed for treatment and a description of which patients were included, excluded, or refused, and why (to be displayed in a flow diagram format, when appropriate).** Example: “A total of 2411 patients were enrolled in the study, 1353 in the HIFU group and 1058 in the surgery group, of whom 586 had myomectomy and 472 hysterectomy (Figure 1). Of 586 myomectomies, 284 (48%) were performed laparoscopically, 233 (40%) by open surgery, and 69 (12%) either transvaginally, by hysteroscopy, or by laparoscopic transvaginal surgery. Of 472 hysterectomies, 251 (53%) were by laparotomy, 93 (20%) by laparoscopy, and 128 (27%) transvaginally. A comparison of baseline data of the two groups is shown in Table 1.”11 |
| Learning Curves | 10 | **Report of learning curve assessment results for operators/team based on pre-defined objective quality metrics, including statistical analysis, if feasible.** Example: “The mean NPV score within the study was 87.2%, with a range for individual centre means from 69.8 to 96.6%. Nineteen of the 20 centres achieved a median NPV higher than the target of 70% set *a priori*. Three centres recorded a mean NPV below the 95% centile, with a 95% CI of 82.5–88.1% (Figure S1). Learning‐curve effects were evaluated using the LC‐CUSUM method, which tests for the presence of a stable ‘in control’ level of performance…. Four centres did not demonstrate stable satisfactory performance by the end of the study. Among the 16 centres that did achieve this, the median number of cases at which stable satisfactory performance was reached was 11 procedures.” 11 |
| Outcomes  | 11 | **Describe results of each pre-specified outcome measure, including patient reported outcome measures, where appropriate. Report the results of pre-specified subgroup analysis to investigate outcomes in patient groups or technique/device variants where pre-study investigator consensus about eligibility was not reached. Report on effects of known confounders.** Example: “Details of clinical experience are shown in Table 4. An access sheath was used in most patients (n = 72). In nine patients with narrow ureters (six children, three females), the ureterorenoscope was inserted directly following rigid ureteroscopy and placement of a guide wire. Mean docking time of the robot was 59.6 s (range: 35–124), which decreased after 42 cases to a mean of 45.9 s. Mean stone location time, including complete inspection of the renal collecting system, was 3.7 min (range: 2–8). Mean frag- mentation time was 46 min (range: 15–118), corresponding to a mean fragmentation speed of 29.1 mm3/min (range: 18–46), increasing to 32.7 mm3/min after 42 cases.”13 |
| Harms | 12 | **Transparent account of all harms or unintended effects reported.** Example: “Minor adverse events occurred in 335 (25%) patients in the HIFU group and in 719 patients (68%) in the surgery group (*P* < 0.001). The events recorded included pain, weakness, or numbness in the lower limbs, back or perineum, haematuria, and general symptoms such as nausea and dizziness. The only categories where HIFU treatment had a higher percentage of minor complications than did surgery were superficial skin burns (2 versus 0), blurred vision and transient pain, and weakness or numbness in the back, shoulder, or lower limb (Table 2). Major adverse events attributable to the intervention occurred in three (0.22%) HIFU cases and in 133 (12.6%) surgical cases within 30 days after treatment. All three HIFU events were second‐degree skin burns. Events in the surgery group included haemorrhage, infection, thromboembolic events, and injury to the bladder….”11 |
| Stakeholder Values | 13 | **Report findings of attempts to evaluate patient and surgeon preferences and values relevant to future RCT trial design and conduct, including any qualitative work done to ascertain views about randomization.** Example: Surgeon engagement will be used to assess the feasibility of using mixed-methods to explore learning and intervention stability in the study. At study entry, the number of surgeons completing the surgeon questionnaire; the number of surgeons consenting to be contacted for interview and the number of surgeons interviewed will be compared. For modifications/learning arising from complications, the numbers of events (modifications and complications) reported will be compared with numbers of free-text responses entered; the number of surgeons consenting to interview and the actual number interviewed.”10 |
| **Discussion** |  |  |
| Limitations | 14 | **Study limitations, addressing sources of potential bias.** Example: **“**Strengths and limitations of the study:In this study HIFU treatment performed very well in terms of both frequency of complications and postoperative recovery period. The comparator surgical treatment groups were not randomised and the surgical treatments were not standardised, exposing comparisons to major risks of bias, but it seems unlikely that the very large differences in some short‐term outcomes could be attributed to selection bias. In China, as in many countries, medical insurance coverage for hospital inpatient stay is determined by the primary diagnosis and treatment. Most Chinese enjoy partial medical insurance coverage with a specific subvention for in‐hospital recovery time, and this may help to explain the relatively long hospital stay in all groups…”11 |
| Stage End-Points | 15a | **Pre-planned consensus review of results, and discussion of appropriateness of progressing to RCT or pilot/feasibility study.** Example:If PPBR is considered safe and data from the screening logs suggests sufficient numbers of PPBR are being performed at participating centres, a consensus meeting with key participating surgeons, steering group members, patients and methodologists will be held to agree the final design for a large-scale pragmatic RCT comparing prepectoral and subpectoral IBBR.”10 |
|  | 15b | **Has agreement been reached about standard technique, including accepted variants, and quality standards based on operator/team experience during this stage?** |
|  | 15c | **Has agreement been reached on the appropriate target patient population and indications, including identification of subgroups for which the applicability of the technique is considered uncertain?**  |
|  | 15d | **Has agreement been reached regarding appropriate outcome measure(s) for a trial, including an estimated power calculation of the primary outcome for a future trial?**  |
|  | 15e | **Has agreement been reached regarding the appropriate comparator treatment for a trial?**  |
|  | 15f | **Are operators and patients willing to accept randomisation between the proposed treatments (establishing equipoise)?**  |
|  | 15g | **Ensure potential harms from learning curves are addressed by evidence of completion, training and mentoring prior to Stage 3, where appropriate.** Example: The learning and modifications emerging from the CRFs and interviews will be consolidated and reviewed at regular intervals by the study team. Any common themes will be shared with participating surgeons using a combination of methods. These will include a ‘tips and tricks’ section on the study website; e-mail updates; study newsletters and potentially social media channels (eg, ‘You Tube’). The acceptability and value of the different approaches to sharing learning will be discussed with individual surgeons at interview and the wider steering group. Metrics such as number of views/hits for online/social media resources will be used to provide additional information…”10 |
| Conclusion | 16 | **Conclusions and relevance, including plans to evaluate the technique/device in a high-quality RCT against the current standard of care. If not planning to further evaluate in IDEAL Stage 3 study, please explain.** Example: “HIFU is now recognised as a fully‐developed state‐of‐the‐art technology, supported by a growing literature, however definitive comparison of the effectiveness and safety of HIFU versus conventional therapies is required before it can be accepted as the standard of care for uterine fibroids….The IDEAL Recommendations propose a pre‐RCT stage (Stage 2b, Exploration) at which potential RCT primary outcomes, operator learning curves, quality control and definitions of the patient group and comparator for an RCT should be worked out, usually in a non‐randomised prospective trial. This IDEAL prospective exploration study has fulfilled all of these objectives, and has demonstrated that HIFU is safe, effective, and deliverable in a large group of unconnected health facilities at high patient volumes. […] Our findings make it clear that RCT proposals comparing complications or short‐term recovery in this context might not attract support or funding, as the very large differences we found are likely to affect equipoise.”11 |
|  |  |  |
| **STAGE 4:** |  |  |
| Title and Abstract | 1a | **Identify the study design and the IDEAL stage in the title or abstract** |
|  | 1b | **Provide a structured summary of background, methods, results, and conclusions** |
| **Introduction** |  |  |
| Background and objectives | 2a | **Identify the study design and the IDEAL stage in the title or abstract.** |
|  | 2b | **Specific objectives stated, which may include recognizing late or uncommon safety outcomes, identification of changes in the use of procedure/device, risk adjustment, quality assurance, and effectiveness measurement.** Example: “We assessed the patterns of changing indications, clinical characteristics, procedural details, complications, and outcomes from the UK TAVI registry.”14 |
| **Methods** |  |  |
| Design | 3 | **Describe the study design (e.g. registry, cohort analysis of real-world data, etc).** Example:“Data were obtained from the Australian Orthopaedic Organisation National Joint Replacement Registry (AOANJRR), collected from September 1, 1999, until December 31, 2013, for 2 cohorts: those who received a cementedGenesis-II cruciate-retaining TKA with a CoCr femoral component and those who received the same design of total knee replacement but with an Oxinium femoral component.”15 |
| Data Source | 4a | **Describe the dataset, including who designed and funds it and who curates and manages it, addressing possible conflicts of interest. Describe whether the dataset was purposely designed for research or collected as a routine database.** Example:“The UK TAVI data set was collected using the Web-based interface from the National Institute of Cardiovascular Outcomes Research. The data set remained unchanged until the end of December 2012.”14 |
|  | 4b | **Structured data fields described or referenced and each item defined, definitions included as an appendix or supplement when necessary.** Example:“The National Cardiovascular Data Repository (NCDR) CathPCI Registry was used as the data source.”16 |
| Participants | 5 | **Detailed account of inclusion and exclusion criteria for subjects.** Example: “Only patients with a valve implanted in the aortic position were included. Patient eligibility for a procedure was decided in each center by a multidisciplinary team composed of interventional cardiologists, imaging cardiologists, cardiothoracic surgeons, and anesthetists. Case ascertainment was performed by comparing the center’s reported numbers of total procedures with the number of procedures uploaded to the National Institute of Cardiovascular Outcomes Research servers.”14 |
| Intervention | 6 | **Describe or reference the intervention/device being monitored, and comparator if applicable.** Example: “The main valve technologies available were the Edwards Sapien and later the Sapien XT, as well as the Medtronic CoreValve. A transfemoral approach was the default strategy for all patients. For those treated with the Medtronic CoreValve, there was later experience with a subclavian and direct aortic approach. For the Edwards Sapien devices, a transapical approach was used when a transfemoral route was not possible, with more recent experience also using a direct aortic approach. From 2011 onward, a small number of valves from St. Jude Medical (Portico), Direct Flow, and JenaValve implanted.”14 |
| Outcomes | 7 | **Description of pre-specified primary and secondary outcome measures selected and how they will be assessed, including patient reported outcome measures, when appropriate, utilising standardised and validated core outcomes sets, when available and applicable. When these are not available, provide rationale for the outcome measure used.** Example: “The primary safety outcome was overall vascular complications, a composite endpoint that included access-site bleeding requiring treatment, access-site hematoma requiring treatment, retroperitoneal bleeding, or any vascular complication requiring intervention. Secondary safety endpoints included access-site bleeding requiring treatment alone and post-procedural blood transfusion. All endpoints and covariates were defined according to the CathPCI version 4 data element definitions. All outcomes were assessed through time of hospital discharge, as limited by the CathPCI Registry dataset.” 16 |
| Statistical methods | 8a | **Statistical methods used to describe and assess patient characteristics and outcomes, including methods of additional analysis (e.g. pre-specified subgroup analysis), and methods to minimize confounding, where appropriate.** Example: “Categorical data are presented as percentages, and comparison between groups was performed by the χ2 test or the Fisher exact test. Numerical data are presented as mean (SD) or median and comparisons were performed with the 2-sample t test or the 2-sample Wilcoxon rank-sum (Mann-Whitney) test. To assess whether there was any trend in the occurrence of complications by the year of operation, a χ2 test for trend was performed, and univariate logistic regression was used to assess the risk of complications... Time-to-event data analysis was performed with the Cox proportional hazards model.…”14 |
|  | 8b | **Explanation of how missing data will be addressed.** Example: “Missing data were imputed using univariate rules, by assuming the absence of a clinical condition for dichotomous variables, and by using the median observed value for continuous variables. Additional sensitivity analyses included alternative approaches to handling missing data, as well as an analysis using center-level matching...”16 |
| **Results** |  |  |
| Baseline data | 9a | **Patient demographic and clinical characteristics, and hospital/centre/operator characteristics.** Example: “There were no differences in age, sex, renal dysfunction, diabetes mellitus, previous myocardial infarction, previous cardiac surgery, previous percutaneous coronary intervention, pulmonary disease, and atrial fibrillation in each successive year. However, there was an increase in the treatment of patients with impaired left ventricular function; the percentage with left ventricular ejection fraction <30% increased from 5.9% in 2007 and 2008 to 10.2% in 2012. In addition, fewer patients with previous stroke and peripheral vascular disease underwent TAVI in 2012 compared with 2007 and 2008.”14 |
|  | 9b | **Indicate extent of missing data for each variable of interest, if appropriate.** Example: “Missing data did not differ between the Mynx-treated patients and those receiving alternative vascular closure devices and represented 0.04% of all risk factor data (Appendix Table S3).”16 |
| Outcomes  | 10 | **Report of main pre-specified outcome measures, including outcome variations among pre-specified subgroups and adjustments for confounders, when applicable** Example: “The absolute risk of a vascular complication was 1.21% with use of the Mynx versus 0.76% with use of an alternate vascular closure device, with a relative risk of 1.59 (95% CI: 1.42-1.78, p<0.001). Each component endpoint of the primary safety outcome was significantly more common in the Mynx-treated patients, including the risk of significant access-site bleeding, vascular complication requiring intervention, significant access-site hematoma, and development of post-procedural retroperitoneal bleeding. Alerts for the secondary endpoints of significant access-site bleeding and post-procedure blood transfusion were triggered after 30 and 15 months, respectively, and were sustained through the remainder of the surveillance period. Protocol-specified propensity-matched analyses were performed for high-risk patient subsets including those age 70 years or greater, females, and diabetics…. 16 |
| **Discussion** |  |  |
| Limitations | 11 | **Discuss limitations, addressing sources of potential bias, known confounders, missing data, and secular trends, including a discussion of the implications of using data that was not originally collected to answer the specific research question, where appropriate.** Example: “Nevertheless, the data are self-reported. Although there are internal consistency and range checks and units are contacted to clarify missing and unexpected values, the data are not subject to independent external validation. Mortality tracking is extremely reliable and independently adjudicated but was available only for England and Wales (which represented the sites of 92% of procedures during this study period). This study is a description of trends and associations, and no conclusions on possible cause and effect can be drawn.”14 |
| Conclusion | 12 | **Conclusions and relevance, including a discussion of the overall interpretation of results, the external validity, and any implications for policy and practice.** Example: “We present the early evolution of TAVI in the United Kingdom using a national registry that includes every patient treated by TAVI and followed up for >6 years. Over this period of time, the overall Logistic EuroSCORE has remained unchanged, although more patients with lower left ventricular ejection fraction were treated. After an initial improvement in survival after 2007 and 2008, short- and long-term outcomes have remained stable, improving further in the most recent (2012) cohort. In addition, a higher percentage of patients were discharged early after the procedure. Short-term (30 day) outcome is difficult to predict on the basis of preprocedural factors, but in the medium and longer term, poor outcome is associated with intrinsic patient characteristics.”14) |
| **Other information** |  |  |
| Protocol | 13 | **Please quote reference or DOI if a protocol was written in advance and made available. If a protocol was not made available, consider including as a supplement if the journal allows.** |
| Ethics | 14 | **Appropriate ethical approvals obtained and informed consent process described. If consent for future use of data was not obtained, please explain.** |
| Funding | 15 | **Sources of funding and support, role of funders, and other conflicts of interest.** |
| Regulatory Approvals | 16 | **Regulatory approvals being sought or obtained (e.g. CE Marking, FDA approval, etc) including the date of approval, if applicable.** |

1. Menon M, Abaza R, Sood A, Ahlawat R, et al. Robotic kidney transplantation with regional hypothermia: Evolution of a novel procedure utilizing the IDEAL guidelines (IDEAL phase 0 and 1). Eur Urol [Internet]. 2014;65(5):1001–1009.

2. Merwe A Van Der, Graewe F, Zühlke A, Barsdorf NW, et al. Penile allotransplantation for penis amputation following ritual circumcision : a case report with 24 months of follow-up. Lancet. 2017;390:1038-1047.

3. Lima E, Rodrigues PL, Mota P, et al. Ureteroscopy-assisted Percutaneous Kidney Access Made Easy: First Clinical Experience with a Novel Navigation System Using Electromagnetic Guidance (IDEAL Stage 1). Eur Urol. 2017 Oct;72(4):610–616.

4. Versteeg AL, van der Velden JM, Hes J, et al. Stereotactic Radiotherapy Followed by Surgical Stabilization Within 24 h for Unstable Spinal Metastases; A Stage I/IIa Study According to the IDEAL Framework. Front Oncol [Internet]. 2018 Dec 20;8(December):1–7.

5. Elsanousi OM, Mohamed MA, Salim FH, et al. Selective devascularization treatment for large hepatocellular carcinoma: Stage 2A IDEAL prospective case series. Int J Surg [Internet]. 2019 Aug;68(April):134–141.

6. Chan JYK, Tsang RK, Holsinger FC, et al. Prospective clinical trial to evaluate safety and feasibility of using a single port flexible robotic system for transoral head and neck surgery. Oral Oncol [Internet]. 2019 Jul;94:101–105.

7. Brännström M, Johannesson L, Dahm-Kähler P, et al. First clinical uterus transplantation trial: a six-month report. Fertil Steril [Internet]. 2014 May;101(5):1228–1236.

8. Diez del Val I, Loureiro C, McCulloch P. The IDEAL prospective development study format for reporting surgical innovations. An illustrative case study of robotic oesophagectomy. Int J Surg [Internet]. 2015 Jul;19:104–111.

9. Blazeby JM, Blencowe NS, Titcomb DR, et al. Demonstration of the IDEAL recommendations for evaluating and reporting surgical innovation in minimally invasive oesophagectomy. Br J Surg [Internet]. 2011 Apr;98(4):544–551.

10. Harvey KL, Mills N, White P, et al. The Pre-BRA (pre-pectoral Breast Reconstruction EvAluation) feasibility study: protocol for a mixed-methods IDEAL 2a/2b prospective cohort study to determine the safety and effectiveness of prepectoral implant-based breast reconstruction. BMJ Open [Internet]. 2020 Jan 26;10(1):e033641.

11. Chen J, Li Y, Wang Z, et al. Evaluation of high-intensity focused ultrasound ablation for uterine fibroids: an IDEAL prospective exploration study. BJOG An Int J Obstet Gynaecol [Internet]. 2018 Feb;125(3):354–364.

12. Buijs M, van Lienden KP, Wagstaff PG, et al. Irreversible Electroporation for the Ablation of Renal Cell Carcinoma: A Prospective, Human, In Vivo Study Protocol (IDEAL Phase 2b). JMIR Res Protoc [Internet]. 2017 Feb 16;6(2):e21.

13. Saglam R, Muslumanoglu AY, Tokatlı Z, et al. A New Robot for Flexible Ureteroscopy: Development and Early Clinical Results (IDEAL Stage 1–2b). Eur Urol [Internet]. 2014 Dec;66(6):1092–1100.

14. Ludman PF, Moat N, de Belder MA, et al. Transcatheter Aortic Valve Implantation in the United Kingdom. Circulation [Internet]. 2015 Mar 31;131(13):1181–1190.

15. Vertullo CJ, Lewis PL, Graves S, et al. Twelve-Year Outcomes of an Oxinium Total Knee Replacement Compared with the Same Cobalt-Chromium Design. J Bone Jt Surg [Internet]. 2017 Feb;99(4):275–283.

16. Resnic FS, Majithia A, Marinac-Dabic D, et al. Registry-Based Prospective, Active Surveillance of Medical-Device Safety. N Engl J Med [Internet]. 2017 Feb 9;376(6):526–535.