Annals of Surgery Publish Ahead of Print
DOI:10.1097/SLA.0000000000006759

Novel Benchmark for Robotic Liver Resection – Bridging Tradition with Innovation

Zhihao Li, MD[1], Matthias Pfister, MD[2], Dimitri A Raptis, MD, MSc, PhD [3],
Yong Ma, MD, PhD[4], Guangchao Yang, MD, PhD[4], Linqiang Li, MD, PhD[4], Xuan Song, MD, PhD[4],
Guillaume Millet, MD[5], Stéphanie Truant, MD, PhD[5], Raphael Venezia, MD[6], Heithem Jeddou, MD[7], Bastien Le Floch, MD[7], Alain Valverde, MD[8], Nicolas Peru, MD[8], Patrick Pessaux, MD, PhD[9],
Fabio Giannone Codiglione, MD, PhD[9,34], Jean-Yves Mabrut, MD, PhD[10], Kayvan Mohkam, MD, PhD[10],
Antonio Sa Cunha, MD, PhD[11], Chady Salloum, MD[11], Pierre Y Blanc, MD, PhD[12],
Bertrand Le Roy, MD, PhD[12], Claire Goumard, MD, PhD[13], Olivier Scatton, MD, PhD[13],
Stylianos Tzedakis, MD, PhD[14], David Fuks, MD, PhD[14], Ismael Chaoui, MD[16], Ashwin Rammohan, MD[17], Silvia Gomes da Silva, MD[19], Mafalda SN Sobral, MD[19], Victor Lopez-Lopez, MD[23],
Leonardo Centonze, MD[24,25], Zhu Zhu, MD, PhD[26], Francesca Ratti, MD, PhD[27], Rebecca Marino, MD[27], Gianluca Rompianesi, MD[28], Yawen Dong, MD[30], Jimmy Shah, MHA[31], Philip C Müller, MD[32],
Taiga Wakabayashi, MD[34], Fabricio F Coelho, MD, PhD[35], Samy Castillo-Flores, MD[41],
Boram Lee, MD[42], Virginia Viti, MD[45], Yutaro Kato, MD[44], Ugo Boggi, MD, PhD[43],
Ho-Seong Han, MD, PhD[42], Patricio Polanco, MD[41], Brian KP Goh, MD[39,40], Tan-To Cheung, MD[38],

Mikhail Efanov, MD, PhD[37], Fabrizio Panaro, MD[36], Marcel Machado, MD[35], Paulo Herman, MD[35],
Go Wakabayashi, MD, PhD[34], Benedetto Ielpo, MD[33], Beat P Müller, MD[32], Alejandro Mejia, MD[31],
Patrick Starlinger, MD, PhD[30], John Martinie, MD[29], Roberto I Troisi, MSc, MD, PhD[28],
Luca Aldrighetti, MD, PhD[27], Guodong Chen, MD, PhD[26], Ricardo Robles-Campos, MD[23],
Gi Hong Choi, MD, PhD[22], Rong Liu, MD, PhD[21], Fabrizio Di Benedetto, MD, PhD[20],
Hugo Pinto Marques, MD[19], Iswanto Sucandy, MD[18], Mohamed Rela, MD[17],
Mathieu D'Hondt, MD, PhD[16], Raffaele Brustia, MD, PhD[6], Olivier Soubrane, MD, PhD[15],
Dieter C Broering, MD, PhD[3], Pierre-Alain Clavien, MD, PhD[1,2]

[1] Department of Surgery and Transplantation, University of Zurich, Zurich, Switzerland.
[2] Wyss Zurich Translational Center, ETH Zurich and University of Zurich, Zurich, Switzerland
[3] King Faisal Specialist Hospital & Research Centre, Riyadh, Saudi Arabia
[4] Department of Minimally Invasive Hepatic Surgery, The First Affiliated Hospital of Harbin Medical University
[5] Digestive Surgery and Transplantation Department, CHU de Lille, University of Lille, 59000, Lille, France
[6] Université Paris-Est Créteil, UPEC, Créteil, France; INSERM, U955, Team "Virus Hépatologie Cancer", Créteil, France; Department of Digestive and Hepatobiliary Surgery, Assistance Publique-Hôpitaux de Paris, Henri Mondor University Hospital, 94000 Créteil, France
[7] Department of Hepatobiliary and Digestive Surgery, Pontchaillou University Hospital, Rennes 1 University, 2, rue Henri-Le-Guilloux, 35000 Rennes, France
[8] Department of Digestive Surgery, Groupe Hospitalier Diaconesses Croix Saint Simon, Paris, France
[9] Department of Digestive Surgery, Strasbourg University Hospital, 67000, Strasbourg, France
[10] Department of Hepatobiliopancreatic Surgery and Liver Transplantation, Hospital Croix Rousse, Lyon, France
[11] Hepato-Biliary Centre, Paul Brousse, AP-HP, Villejuif, Paris, France
[12] Department of Visceral and Digestive Surgery, Centre Hospitalier Universitaire de Saint-Etienne, Saint-Etienne, France
[13] Department of Hepatobiliary Surgery and Liver Transplantation, Sorbonne University, Hôpital Pitié-Salpêtrière, Assistance Publique- Hôpitaux de Paris, France
[14] Department of Hepatobiliary, Digestive and Endocrine Surgery, Université Paris Cité, Cochin Hospital, AP-HP, Paris, France
[15] Department of Digestive, Oncological and Metabolic Surgery, Institut Mutualiste Montsouris, Paris, France
[16] Department of Digestive and Hepatobiliary/Pancreatic Surgery, Groeninge Hospital, Kortrijk, Belgium
[17] Dr Rela Institute and Medical Center, Chennai, India

[18] Digestive Health Institute at AdventHealth Tampa, Florida USA
[19] Hepato-Biliary-Pancreatic and Transplantation Centre, Centro Hospitalar Universitário de Lisboa Central, Lisbon, Portugal
[20] Hepato-Pancreato-Biliary Surgery and Liver Transplantation Unit, University of Modena and Reggio Emilia, Modena, Italy
[21] Faculty of Hepatopancreatobiliary Surgery, the First Medical Center of Chinese People's Liberation Army (PLA) General Hospital, Beijing, China
[22] Division of Hepatopancreatobiliary Surgery, Department of Surgery, Severance, Hospital, Yonsei University College of Medicine, Seoul, Korea
[23] Department of General, Visceral and Transplantation Surgery, Clinic and University Hospital Virgen de la Arrixaca, Murcia, Spain
[24] Department of General Surgery and Transplantation, Niguarda Ca' Granda Hospital, Milan, Italy
[25] Clinical and Experimental Medicine PhD Program, University of Modena and Reggio Emilia, Modena, Italy
[26] Department of Hepatopancreatobiliary Surgery, The First Affiliated Hospital, Hengyang Medical School, University of South China, Hengyang, China
[27] Division of Hepatobiliary Surgery, IRCCS San Raffaele Hospital, Vita-Salute University, Milan, Italy
[28] Division of HPB, Minimally Invasive and Robotic Surgery, Federico II University Hospital Naples, Naples, Italy
[29] Division of HPB Surgery, Carolinas Medical Center, Atrium Health, Charlotte, USA
[30] Department of Hepatobiliary and Pancreas Surgery, Mayo Clinic Rochester, Rochester, USA
[31] The Liver Institute, Methodist Dallas Medical Center, Dallas, USA
[32] Clarunis University Centre for Gastrointestinal and Liver Diseases, Basel, Switzerland; Department of Visceral Surgery, University Hospital Basel, Basel, Switzerland
[33] HPB Unit, Hospital del mar, Barcelona, Spain
[34] Center for Advanced Treatment of Hepatobiliary and Pancreatic Diseases, Ageo Central General Hospital, Saitama, Japan
[35] Department of Surgery, Hospital das Clinicas, University of São Paulo Medical School, Brazil.
[36] SCDU General Surgery, Surgical Oncology, Minimally Invasive, Robotic and HBP Surgery, AOUAL SS. Antonio e Biagio e Cesare Arrigo, Alessandria, Italy
[37] Department of Hepato-Pancreato-Biliary Surgery, Moscow Clinical Scientific Center, Moscow, Russia
[38] Department of Surgery, Queen Mary Hospital, The University of Hong Kong, Hong Kong SAR, China
[39] Department of Hepatopancreatobiliary and Transplant Surgery, Singapore General Hospital, National Cancer Centre Singapore
[40] Surgery, Academic Clinical Program, Duke-National Medical School, Singapore
[41] Department of Surgery, University of Texas Southwestern Medical Center, Dallas, USA
[42] Department of Surgery, Seoul National University Hospital Bundang, Seoul National University College of Medicine, Seoul, Korea
[43] Division of General and Transplant Surgery, University of Pisa, Pisa, Italy
[44] Department of Gastroenterological Surgery, Bantane Hospital, Fujita Health University, Nagoya, Japan

✉**Corresponding Author:**
Pierre-Alain Clavien, MD, PhD
Professor of Surgery & Transplantation,
University of Zurich
Email: clavien@access.uzh.ch
Manuscript word count: 2639
Tables: 3, Figures: 2; Supplementary Tables: 2, Figures: 1

Short Running Head: Benchmark for Robotic Liver Resection

Keywords: Robotic liver resection, Benchmarking, Morbidity, Complication, Iwate score, Textbook outcome, Patient selection.

ABSTRACT

Objective: To establish benchmark cutoffs for robotic liver resection (R-LR), encompassing both major and minor resections, and to determine the impact of patient selection on outcomes.

Background: R-LR is a key advancement in minimally invasive liver surgery but lacks standardized benchmarks, especially for minor resections. While guidelines endorse R-LR, its role in optimizing outcomes remains unclear. This study establishes the first benchmarks for R-LR, enabling comparisons across surgical modalities and refining patient selection.

Methods: This retrospective, multicenter study analyzed consecutive adult patients undergoing R-LR at 30 international centers (2020–2023). Benchmark centers had an annual case volume of ≥15 R-LR. Benchmark criteria included ASA ≤2, no major comorbidities, no prior liver resections, and Child-Pugh A status. Benchmark cutoffs for 14 key outcomes were set at the 50th and 75th percentiles of median values across benchmark centers. Multivariable logistic regression identified predictors of textbook outcomes.

Results: Eighteen high-volume centers contributed 4028 cases with 2632 (65.3%) meeting benchmark criteria. Malignancy was the indication in 29.6%, most commonly hepatocellular carcinoma followed by colorectal liver metastases. Major liver resection was performed in 42.6%. The distribution of Iwate difficulty scores was low (25.4%), intermediate (53.9%),

advanced (15.9%), and expert (4.9%). Benchmark cutoffs were established for minor and major resection, and stratified by Iwate low (0-3), intermediate (4-6) and high (7-12): Open conversion (≤6.5% minor, ≤10.5% major), major complications within 90 days (≤5.2% minor, ≤16.7% major), R1 resection (≤9.2% minor, ≤6.7% major). Benchmark cases performed at low-volume centers were able to achieve outcomes within the corresponding benchmark cutoffs. In fact, patient selection reflected by the proportion of benchmark patients, rather than case volume, was associated with textbook outcomes.

Conclusions: This study defines R-LR benchmarks, emphasizing patient selection over center volume for optimal outcomes. Benchmark cutoffs guide training and support the expansion of R-LR.

INTRODUCTION

Minimally invasive liver surgery (MILS) has reached the assessment phase (stage 3) of the IDEAL framework for the development of surgical innovations[1,2]. While robotic liver resection (R-LR) represents a key technological advancement, its adoption has lagged behind laparoscopic approaches, mostly due to cost-effectiveness concerns. Despite the establishment of benchmarks for open[3,4] and laparoscopic liver surgery[5], standardized benchmarks for R-LR remain absent. Existing benchmark studies also focus almost exclusively on major resections, overlooking minor resections, which constitute a significant proportion of MILS. This omission is particularly relevant given that the Southampton Consensus Guidelines[6] recognize MILS as the standard approach for minor liver resections, and the Paris Consensus on robotic Hepato-Pancreato-Biliary Surgery[7] strongly endorses robotic anatomical and non-anatomical minor resections as valid alternatives to laparoscopy and recommends establishing benchmarks for robotic procedures. Additionally, R-LR has gained increasing importance in donor hepatectomy[8], with several centers demonstrating excellent outcomes, yet this application has not been addressed in existing benchmark studies. While robotic surgery brings significant technological innovation to liver surgery, the role of patient selection in optimizing outcomes remains unclear, raising the question of when and where R-LR provides the greatest benefit.

To address these gaps, this international, high-volume study establishes the first benchmarks for R-LR, encompassing both major and minor resections. By analyzing patient characteristics and surgical complexity alongside outcomes, we define the impact of patient selection on R-LR success. This study builds upon prior benchmark research in open and laparoscopic liver surgery, enabling direct comparisons across surgical modalities and guiding the future evolution of R-LR.

METHODS

Study Design

This study included consecutive adult patients (≥18 years) who underwent anatomical and non-anatomical R-LR between January 2020 and December 2023. Patients who underwent R-LR with concurrent major procedures (e.g., robotic colectomies) were excluded. Benchmarks were established using standardized methodology[9], as previously applied to other complex

procedures[3,10,4,11]. High-volume benchmark centers were defined by: (a) performing ≥15 R-LRs per year or ≥60 over 4 years, (b) prior publications on R-LR, and (c) maintaining a prospective patient database with at least one year of follow-up. Low-volume centers (≤15 cases per year) that did not qualify for benchmark analysis were included as a control group. The benchmark cohort consisted of patients meeting the following criteria: American Society of Anesthesiology (ASA) classification ≤2, absence of major comorbidities (recent myocardial infarction, advanced renal or obstructive pulmonary disease), no prior liver resections, and Child-Pugh A liver function. These criteria aligned with previous benchmark analyses for liver resections[4,11], ensuring comparability among benchmark cohorts. Patients who did not meet benchmark criteria formed the non-benchmark cohort, serving as another control group. The primary endpoints of this study are the benchmark values outlined in the next section. Ethical approval was obtained from the Cantonal Ethics Commission of Zurich (BASEC 2024-01513) and each participating center.

Definition of Benchmark Values and Other Variables

Fourteen clinically relevant outcome parameters were identified for benchmark establishment, aligning with established liver resection benchmark studies[3,4,11]. Intraoperative Benchmarks: operation time, estimated blood loss, intraoperative blood transfusion, and open conversion. Postoperative Benchmarks: 90-day morbidity, major morbidity (≥ Clavien-Dindo grade IIIa[12]), mortality, reoperation, unplanned readmission; postoperative liver failure and bile leak (ISGLS grade B/C[13,14]); length of stay; and R1 resection (<1 mm margin for malignant tumors). Textbook outcome was defined as the absence of intraoperative incidents (Oslo classification grade ≥2)[15], 90-day major morbidity or mortality, postoperative bile leak or liver failure (ISGLS grade B/C), and R1 resection[16]. Resection extent was classified according to the 2000 Brisbane classification[17] and New World terminology[18]. Surgical difficulty was graded using the Iwate score[19].

Statistical Analysis

Descriptive statistics were presented as medians with interquartile ranges (IQR) for continuous variables and counts with percentages for categorical variables. Continuous variables were compared using the Mann-Whitney U test, while categorical variables were analyzed with

the Chi-square test. Benchmark cutoffs were defined as the 75th percentile of the median values across all benchmark centers for each negative outcome indicator and the 25th percentile for positive outcome indicators, such as textbook outcome. We performed a sensitivity analysis, excluding donor hepatectomies, to determine whether this exclusion significantly altered the benchmark cutoffs.

The Pearson correlation coefficient (R^2) was calculated to evaluate the association between center characteristics and outcomes. Multivariable logistic regression with backward stepwise selection was used to identify predictors of textbook outcome. Results were reported as odds ratios (OR) with 95% confidence intervals (CI). We investigated the relationship between the Iwate score and the odds of *not* achieving a textbook outcome, stratified by benchmark patient status. A logistic regression model was constructed, incorporating an interaction term between the Iwate score and benchmark patient status to assess how the effect of the Iwate score varies between benchmark and non-benchmark patients. To visualize the potentially non-linear effect of the Iwate score on the OR (comparing benchmark patients to non-benchmark patients), a smoothing spline with 4 degrees of freedom was fitted to the calculated ORs and their 95% CI limits. Statistical analyses were conducted using R 4.4.1 (R Foundation for Statistical Computing, Vienna, Austria).

RESULTS

Benchmark Center and Case Characteristics

A total of 30 centers across four continents (Asia (N=12), Europe (N=12), North America (N=5), and South America (N=1)) participated in this study, contributing 4,369 cases. Among them, 18 centers met the criteria for high-volume benchmark centers, accounting for 4,028 cases (92.2% of the total) (Supplementary Figure 1, Supplemental Digital Content 1, http://links.lww.com/SLA/F486). Of these, 2,632 cases (65.3%) qualified as benchmark cases. Patient characteristics for the overall population and the high-volume benchmark centers are summarized in Supplementary Tables 1 and 2, Supplemental Digital Content 1, http://links.lww.com/SLA/F486, respectively. Among the 2632 benchmark cases used to establish benchmark cutoffs, the median age was 39.9 years (IQR: 29.0-59.0), and 59.6% were male. Liver resection was performed for donor hepatectomy in 46.9% of cases, malignancy in

29.3% (primarily hepatocellular carcinoma, followed by colorectal liver metastases), and benign lesions in 22.8%. Major liver resection was performed in 42.6% of cases, while anatomical resections accounted for 71.3%. The distribution of Iwate difficulty scores was low (25.4%), intermediate (53.9%), advanced (15.9%), and expert (4.9%) (Supplementary Table 2, Supplemental Digital Content 1, http://links.lww.com/SLA/F486).

Benchmark Cutoffs

Separate benchmark cutoffs were established for five categories of R-LR: Minor and major resection, Iwate low (0-3), intermediate (4-6) and high (7-12) liver resections (Table 1). Open conversion rates were ≤6.5% for minor resections, and ≤10.5% for major resections. The cutoffs for major complications (Clavien-Dindo ≥3a) within 90 days were ≤5.2% for minor resections, and ≤16.7% (P75) for major resections. The 90-day mortality cutoff was 0% across all categories. The cutoffs for bile leak (ISGLS Grade B/C) were ≤2.6% for minor resections, and ≤6.7% for major resections. The benchmark cutoffs for textbook outcome were ≥82.0% for minor resections, and ≥68.4% for major resections. The cutoffs for R1 resection were ≤9.2% for minor resections, and ≤6.7% for major resections. Benchmark cutoffs increased from low to high Iwate scores, except for Textbook Outcome, which showed an inverse trend. The benchmark cutoffs from the sensitivity analysis, excluding donor hepatectomies, are presented in Supplementary Table S3, Supplemental Digital Content 1, http://links.lww.com/SLA/F486. Despite slight increases in the cutoffs for open conversion rate, major complications, and R1 resection rate, these changes were minimal and did not impact the overall findings.

Comparison of Benchmark Cutoffs with Non-Benchmark Cases and Low-Volume Centers

Non-benchmark cases performed at benchmark centers exhibited outcomes exceeding the 75th percentile benchmark cutoff for several key metrics: any complication rate, re-laparotomy, readmission, mortality, liver failure, bile leak, and R1 resection rates (Table 2). While in non-benchmark centers outcomes of minor resections were within benchmark cutoffs, major resections showed increased operation time, blood loss, open conversion rate, major complications, mortality, liver failure, bile leak, and R1 resection rate. It is important to acknowledge that this control group (n=27) is small, potentially increasing the risk of type II error. When stratified by the three Iwate difficulty classes, low-volume centers were able to

achieve outcomes within the corresponding benchmark cutoffs. Outcomes following donor hepatectomy were highly favorable and consistent with the benchmark cutoffs established for both major and minor liver resections.

Correlation between Center Characteristics and Outcomes

Figure 1 illustrates the relationship between center-specific characteristics—the proportion of benchmark cases and annual case volume—and key postoperative outcomes, including major complications and textbook outcomes, for all participating centers. A higher proportion of benchmark cases within a center was associated with lower rates of major complications and higher textbook outcome rates (Figure 1A&B). However, no significant correlation was observed between annual case volume and either outcome (Figure 1C&D). The median Iwate score of each center, representing the surgical complexity of cases, correlated with the rate of major complications (Figure 1E). Centers with a low proportion of highly complex cases (Iwate 8-12) achieved textbook outcomes in low-complexity cases (Iwate <4) at a level comparable to centers managing more complex cases.

Factors Associated with Not Achieving Textbook Outcome

Multivariable logistic regression identified five independent factors associated with not achieving a textbook outcome (Table 3). Meeting benchmark criteria was the only protective factor (OR: 0.6, 95% CI: 0.48–0.71). In contrast, malignant indication (OR: 5.0, 95% CI: 3.89–6.59), neoadjuvant chemotherapy (OR: 1.5, 95% CI: 1.18–1.87), major hepatectomy (OR: 1.6, 95% CI: 1.29–2.00), and Iwate score (OR: 1.1, 95% CI: 1.03–1.12) were identified as risk factors.

Figure 2 illustrates the association between patient selection based on benchmark criteria and the likelihood of not achieving a textbook outcome across the spectrum of surgical difficulty, represented by the Iwate score. Across all Iwate scores, benchmark patients had a consistently lower risk of not achieving a textbook outcome, with the OR remaining below 1. The protective effect of meeting benchmark criteria was strongest at lower Iwate scores but gradually diminished and became more variable as the Iwate score increased.

DISCUSSION

With the growing adoption of robotic liver surgery, this multicenter study reflects current global expertise and establishes critical benchmarks for these evolving procedures. Employing established benchmarking methodology, we defined cutoffs for key outcome metrics in R-LR, stratified by both the extent of resection (major/minor) and surgical difficulty (Iwate score). Comparison of these novel robotic benchmarks with previously established cutoffs for open major hepatectomy revealed improvements in most outcomes, supporting the transition to a robotic approach. Furthermore, our analysis revealed that patient selection, rather than center case volume, is the primary determinant of achieving textbook outcomes. This finding suggests that low-volume centers can achieve benchmark-level performance through careful patient selection. This study represents the most comprehensive analysis to date of robotic liver surgery, evaluating the interplay of patient selection, surgical complexity, and perioperative outcomes, and informing future directions for this technology.

Stratifying patients according to the same benchmark criteria as in open surgery[4] and laparoscopy[11] offers the significant advantage of enabling comparisons among different surgical approaches. Our analysis reveals that, compared to the benchmark for open major hepatectomies, the benchmark for R-LR exhibited a longer operation time by one hour, but significantly reduced intraoperative blood loss by 500 mL and hospital stay by approximately five days. While prior single-center studies[20–23] have suggested similar trends, this multicenter study provides robust confirmation of these benefits on a global scale, utilizing a matched population that fulfills established benchmark criteria. A remaining question is whether R-LR offers advantages over the open approach in terms of perioperative morbidity and mortality. Our analysis of benchmark criteria provides compelling evidence in favor of robotic surgery. Specifically, cutoffs for major complications, readmission rates, bile leaks, liver failure, mortality, and R1 resection rates were all lower for R-LR than those established for the open approach[4]. This demonstrates a clear perioperative outcome benefit for R-LR within high-volume centers. Extending this comparison to robotic donor hepatectomy, major complication and liver failure rates were significantly below established benchmark cutoffs for open donor hepatectomy. This supports recent findings that robotic surgery offers lower donor morbidity compared to open and laparoscopic approaches[24]. Comparing outcomes to laparoscopic liver resection is less straightforward because benchmark

cutoffs were established for left and right hepatectomies, not strictly for major resections as a whole[11].

Despite the existence of several benchmark studies in liver resection, minor liver resections have been largely overlooked, representing a significant knowledge gap. The frequency and clinical significance of minor liver resections have increased substantially over the past decade. This rise is attributable to several factors, including technological advancements, refined surgical techniques, a growing emphasis on parenchymal preservation[25], and an expanded role for surgical management of colorectal liver metastases[26]. While minor liver resections are generally associated with lower morbidity than major resections, as evidenced by our benchmark cutoffs, this does not preclude significant surgical complexity. Non-anatomical resections often necessitate complex multiplanar dissection trajectories, making it challenging for the surgeon to accurately assess tumor resection margins. This difficulty is reflected in the higher benchmark cutoffs for R1 resections observed in minor resections compared to major resections. Additionally, the anatomical location of lesions and their proximity to major vasculature are significant determinants of surgical complexity. To account for this variability in surgical difficulty, the Iwate scoring system was developed[19,27]. While originally designed for laparoscopic liver resections, the Iwate score has also been validated for R-LRs[28–30]. In this study, we employed the Iwate score to stratify R-LR according to their complexity, irrespective of the conventional major/minor resection classification. Increasing Iwate classes correlated with increasing benchmark cutoffs, highlighting the impact of surgical complexity on outcomes and demonstrating the utility of the Iwate score in providing a more nuanced and accurate application of benchmark criteria.

Benchmark and textbook outcome studies represent distinct, yet complementary, approaches to evaluating surgical performance. While benchmark criteria are traditionally used to define an ideal patient population for establishing optimal outcomes, this study also utilizes the percentage of benchmark cases as a metric of patient selection. Interestingly, and in contrast to findings in laparoscopic liver resection[11,31], center case volume did not correlate with outcomes in this analysis. This finding is further supported by the ability of low-volume centers to achieve benchmark outcomes, particularly for minor resections where sample sizes were

adequate. Furthermore, expertise in highly complex R-LR does not appear to be necessary or provide additional benefits for achieving textbook outcomes in less complex cases. Instead, the proportion of benchmark patients (i.e., patient selection) emerged as a key determinant of success. This finding suggests that the technological advancements inherent in robotic surgery may mitigate the steep learning curve typically associated with laparoscopic liver resection, potentially reducing the influence of case volume and amplifying the importance of patient selection. This interpretation is supported by our analysis, which demonstrates that fulfilling benchmark criteria was a protective factor against failing to achieve textbook outcomes across the Iwate difficulty scores. However, this protective effect diminished with increasing surgical difficulty, indicating that optimal patient selection cannot fully compensate for the challenges posed by more complex procedures.

This study is subject to several limitations. First, as a retrospective study, it is inherently susceptible to biases associated with data collection and reporting that cannot be fully controlled for. Second, the case volume varied considerably among participating centers, with larger centers contributing most cases. This imbalance could potentially skew the overall dataset, introducing bias towards the practices and outcomes of high-volume centers. However, the benchmark methodology mitigates this by evaluating data at the center level, thereby distributing the analytical weight equally across institutions. Third, while this study design allows for the assessment of surgical and perioperative outcomes using the benchmark methodology, it is not suitable for evaluating oncologic outcomes. Benchmark criteria, by design, do not account for specific tumor characteristics (e.g., entity, grading, staging), making direct comparisons of survival across different tumor biologies methodologically unsound. Finally, although robotic donor hepatectomy might be considered a separate procedure, we included it because healthy donors provide an optimal benchmark population, and the fundamental surgical steps are consistent with those of other indications. The results of our sensitivity analysis reinforced this decision.

In conclusion, this study establishes critical benchmark outcomes for R-LR, supporting its safe and effective adoption. Our findings highlight the advantages of the robotic approach over open and potentially laparoscopic resection, with a less steep learning curve that enables

low-volume centers to achieve benchmark outcomes. However, challenges remain, including cost, accessibility, standardized training, and technological advancements such as haptic feedback and AI integration. Despite these innovations, meticulous patient selection remains paramount, with benchmark criteria strongly predicting success. By defining objective performance targets, this study informs training, guides evaluation of emerging technologies, and supports data-driven refinement of robotic liver surgery. Ultimately, balancing innovation with core surgical principles—careful patient selection, precise technique, and continuous improvement—is essential to maximizing patient benefit from robotic liver surgery.

ACKNOWLEDGEMENT

The data from the French centers participating in this study are derived from the FRIES (**F**rench **R**obotic L**i**v**e**r **S**urgery) cohort, conducted under the auspices of the *Association de Chirurgie Hépato-Bilio-Pancréatique et Transplantation* (ACHBPT-FRIES-2024*)*

REFERENCES

1. Ergina PL, Barkun JS, McCulloch P, Cook JA, Altman DG. IDEAL framework for surgical innovation 2: observational studies in the exploration and assessment stages. Published online June 18, 2013. doi:10.1136/bmj.f3011

2. Deyrat J, Fuks D, Murris J, et al. Evolution of minimally invasive liver surgery in France over the last decade. *Surg Endosc*. 2024;38(8):4457-4467. doi:10.1007/s00464-024-10951-3

3. Rössler F, Sapisochin G, Song G, et al. Defining Benchmarks for Major Liver Surgery: A multicenter Analysis of 5202 Living Liver Donors. *Ann Surg*. 2016;264(3):492-500. doi:10.1097/SLA.0000000000001849

4. Sousa Da Silva RX, Breuer E, Shankar S, et al. Novel Benchmark Values for Open Major Anatomic Liver Resection in Non-cirrhotic Patients: A Multicentric Study of 44 International Expert Centers. *Ann Surg*. 2023;278(5):748-755. doi:10.1097/SLA.0000000000006012

5. Goh BKP, Han HS, Chen KH, et al. Defining Global Benchmarks for Laparoscopic Liver Resections: An International Multicenter Study. *Ann Surg*. 2023;277(4):e839-e848. doi:10.1097/SLA.0000000000005530

6. Abu Hilal M, Aldrighetti L, Dagher I, et al. The Southampton Consensus Guidelines for Laparoscopic Liver Surgery: From Indication to Implementation. *Ann Surg*. 2018;268(1):11-18. doi:10.1097/SLA.0000000000002524

7. Hobeika C, Pfister M, Geller D, et al. Recommendations on Robotic Hepato-Pancreato-Biliary Surgery. The Paris Jury-Based Consensus Conference. *Ann Surg*. 2025;281(1):136. doi:10.1097/SLA.0000000000006365

8. Broering DC, Raptis DA, Malago M, Clavien PA, MIOT Collaborative. Revolutionizing Organ Transplantation With Robotic Surgery. *Ann Surg*. 2024;280(5):706-712. doi:10.1097/SLA.0000000000006460

9. Gero D, Muller X, Staiger RD, et al. How to Establish Benchmarks for Surgical Outcomes?: A Checklist Based on an International Expert Delphi Consensus. *Ann Surg*. 2022;275(1):115-120. doi:10.1097/SLA.0000000000003931

10. Li Z, Rammohan A, Gunasekaran V, et al. Novel Benchmark for Adult-to-Adult Living-donor Liver Transplantation: Integrating Eastern and Western Experiences. *Ann Surg*. 2023;278(5):798-806. doi:10.1097/SLA.0000000000006038

11. Goh BKP, Han HS, Chen KH, et al. Defining Global Benchmarks for Laparoscopic Liver Resections: An International Multicenter Study. *Ann Surg*. 2023;277(4):e839-e848. doi:10.1097/SLA.0000000000005530

12. Dindo D, Demartines N, Clavien PA. Classification of Surgical Complications. *Ann Surg*. 2004;240(2):205-213. doi:10.1097/01.sla.0000133083.54934.ae

13. Rahbari NN, Garden OJ, Padbury R, et al. Posthepatectomy liver failure: a definition and grading by the International Study Group of Liver Surgery (ISGLS). *Surgery*. 2011;149(5):713-724. doi:10.1016/j.surg.2010.10.001

14. Koch M, Garden OJ, Padbury R, et al. Bile leakage after hepatobiliary and pancreatic surgery: a definition and grading of severity by the International Study Group of Liver Surgery. *Surgery*. 2011;149(5):680-688. doi:10.1016/j.surg.2010.12.002

15. Kazaryan AM, Røsok BI, Edwin B. Morbidity Assessment in Surgery: Refinement Proposal Based on a Concept of Perioperative Adverse Events. doi:10.1155/2013/625093

16. Görgec B, Benedetti Cacciaguerra A, Pawlik TM, et al. An International Expert Delphi Consensus on Defining Textbook Outcome in Liver Surgery (TOLS). *Ann Surg*. 2023;277(5):821. doi:10.1097/SLA.0000000000005668

17. Strasberg SM, Belghiti J, Clavien PA, et al. The Brisbane 2000 Terminology of Liver Anatomy and Resections. *HPB*. 2000;2(3):333-339. doi:10.1016/S1365-182X(17)30755-4

18. Nagino M, DeMatteo R, Lang H, et al. Proposal of a New Comprehensive Notation for Hepatectomy: The "New World" Terminology. *Ann Surg*. 2021;274(1):1. doi:10.1097/SLA.0000000000004808

19. Wakabayashi G, Cherqui D, Geller DA, et al. Recommendations for laparoscopic liver resection: a report from the second international consensus conference held in Morioka. *Ann Surg*. 2015;261(4):619-629. doi:10.1097/SLA.0000000000001184

20. Kingham TP, Leung U, Kuk D, et al. Robotic Liver Resection: A Case-Matched Comparison. *World J Surg*. 2016;40(6):1. doi:10.1007/s00268-016-3446-9

21. Machairas N, Papaconstantinou D, Tsilimigras DI, et al. Comparison between robotic and open liver resection: a systematic review and meta-analysis of short-term outcomes. *Updat Surg*. 2019;71(1):39-48. doi:10.1007/s13304-019-00629-0

22. Klein B, Holzner ML, Llore N, Radkani P, Winslow E, Hawksworth JS. Open vs Robotic Hepatectomy: Propensity-matched Analysis by ... : Journal of the American College of Surgeons. Accessed March 4, 2025. https://journals.lww.com/journalacs/fulltext/2023/05003/open_vs_robotic_hepatectomy__propensity_matched.168.aspx

23. Yang HY, Rho SY, Han DH, Choi JS, Choi GH. Robotic major liver resections: Surgical outcomes compared with open major liver resections. *Ann Hepato-Biliary-Pancreat Surg*. 2021;25(1):8-17. doi:10.14701/ahbps.2021.25.1.8

24. Broering DC, Elsheikh Y, Alnemary Y, et al. Robotic Versus Open Right Lobe Donor Hepatectomy for Adult Living Donor Liver Transplantation: A Propensity Score-Matched Analysis. *Liver Transplant Off Publ Am Assoc Study Liver Dis Int Liver Transplant Soc*. 2020;26(11):1455-1464. doi:10.1002/lt.25820

25. Torzilli G, McCormack L, Pawlik T. Parenchyma-sparing liver resections. *Int J Surg*. 2020;82:192-197. doi:10.1016/j.ijsu.2020.04.047

26. Clavien PA, Petrowsky H, DeOliveira ML, Graf R. Strategies for Safer Liver Surgery and Partial Liver Transplantation. *N Engl J Med*. 2007;356(15):1545-1559. doi:10.1056/NEJMra065156

27. Tanaka S, Kawaguchi Y, Kubo S, et al. Validation of index-based IWATE criteria as an improved difficulty scoring system for laparoscopic liver resection. *Surgery*. 2019;165(4):731-740. doi:10.1016/j.surg.2018.10.012

28. Luberice K, Sucandy I, Modasi A, et al. Applying IWATE criteria to robotic hepatectomy: is there a "robotic effect"? *HPB*. 2021;23(6):899-906. doi:10.1016/j.hpb.2020.10.008

29. Steinkraus KC, Traub B, Heger P, et al. Results of robotic liver surgery in association with IWATE criteria — the first 100 cases. *Langenbecks Arch Surg*. 2024;409(1):50. doi:10.1007/s00423-024-03239-6

30. Lamberty SA, Hoelzen JP, Katou S, et al. Validation of the IWATE Criteria in Robotic-Assisted Liver Resections. *J Clin Med*. 2024;13(9):2697. doi:10.3390/jcm13092697

31. Russolillo N, Aldrighetti L, Cillo U, et al. Risk-adjusted benchmarks in laparoscopic liver surgery in a national cohort. *Br J Surg*. 2020;107(7):845-853. doi:10.1002/bjs.11404

Figure 1: Linear Correlation Between Center Characteristics and Outcomes.

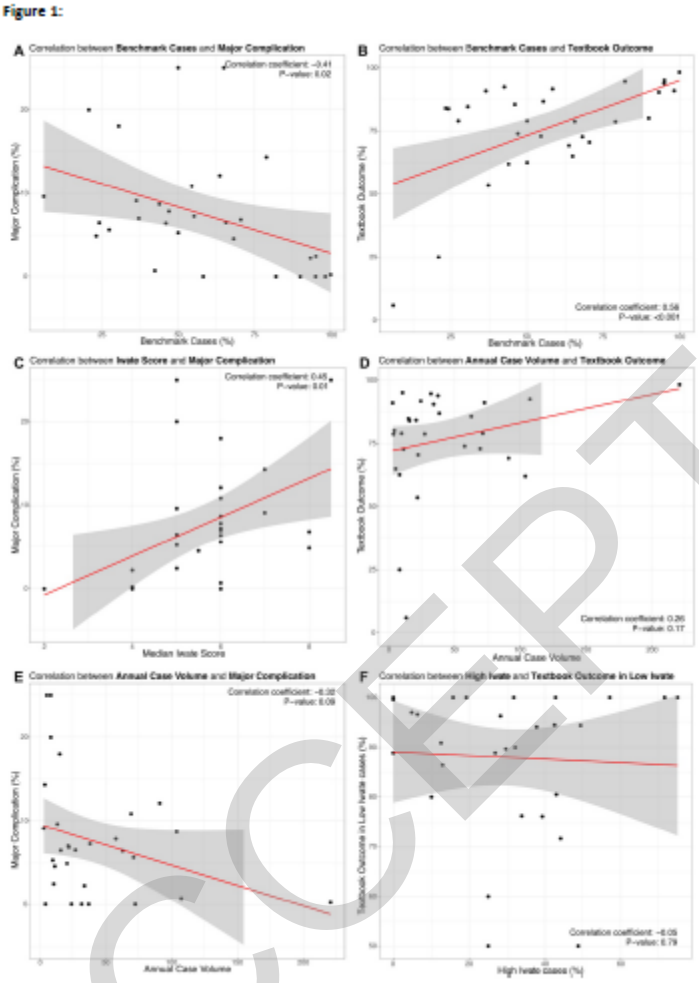

Figure 2: Association of Being a Benchmark Patient on the Risk of Not Achieving Textbook Outcome Across Iwate Scores.

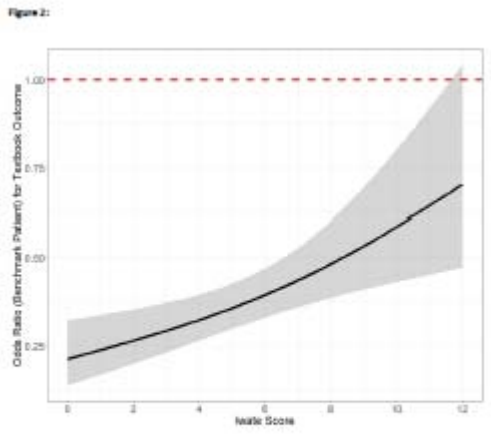

Table 1: Benchmark Cutoffs for Robotic Liver Resections (Major, Minor; Iwate Low, Intermediate, High).

	Minor Resection		Major Resection		Iwate Low (0-3)		Iwate Intermed. (4-6)		Iwate High (7-12)	
Benchmark Cutoffs	**P50**	**BM**	**P50**	**BM**	**P50**	**BM**	**P50**	**BM**	**P50**	**BM**
Perioperative Metrics										
Operation time (min)	≤198	≤294	≤345	≤437	≤179	≤234	≤219	≤354	≤219	≤357
Estimated blood loss (mL)	≤100	≤100	≤200	≤200	≤50	≤83	≤100	≤138	≤100	≤200
Intraoperative blood transfusion (u)	≤0	≤0	≤0	≤0	≤0	≤0	≤0	≤0	≤0	≤0
Open conversion (%)	≤0.2	≤6.5	≤0.7	≤10.5	≤0	≤0	≤0.6	≤3.5	≤0.6	≤9.4
Hospital stay (d)	≤4	≤5	≤5	≤7	≤3	≤5	≤4	≤6	≤4	≤7
Postoperative Morbidity at 90d										
Any complication (%)	≤8.2	≤10.4	≤16.3	≤30.4	≤6.3	≤16.4	≤10.2	≤16.5	≤10.2	≤21.5
Clavien–Dindo grade ≥3a (%)	≤1.8	≤5.2	≤2.6	≤16.7	≤0	≤2.2	≤0.1	≤5.9	≤0.1	≤9.5
CCI®	≤0	≤0	≤0	≤0	≤0	≤0	≤0	≤0	≤0	≤0
Relaparotomy rate (%)	≤0	≤0	≤0	≤0	≤0	≤0	≤0	≤0.2	≤0	≤0
Readmission rate (%)	≤2.2	≤5.5	≤0	≤7.4	≤0	≤4.2	≤1.3	≤5.0	≤1.3	≤5.0
Mortality (%)	≤0	≤0	≤0	≤0	≤0	≤0	≤0	≤0	≤0	≤0
Liver failure – ISGLS Grade B/C (%)	≤0	≤0	≤0	≤0	≤0	≤0	≤0	≤0	≤0	≤0.3
Bile leak – ISGLS Grade B/C (%)	≤1.1	≤2.6	≤1.3	≤6.7	≤0	≤0	≤0	≤1.8	≤0	≤7.1
Textbook outcome (%)	≥93.0	≥82.0	≥80.0	≥68.4	≥94.4	≥84.2	≥89.5	≥76.2	≥83.2	≥67.9
Oncological Outcome										
R1 resection (%)	≤2.2	≤9.2	≤0	≤6.7	≤0	≤5.7	≤0	≤5.7	≤0	≤14.1

P50 represents the 50th and BM the 75th-percentile of centers median.
Abbreviations: BM, Benchmark; CCI®, Comprehensive Complication Index; ISGLS, International Study Group of Liver Surgery

Table 2: Benchmark Outcome Measures Across Three Control Groups (High-risk Non-Benchmark Patients at Expert Centers, Low-risk Benchmark Patients at Non-Expert Centers, and Living Donor Hepatectomy Cohort).

	Minor Resection		Major Resection		Iwate Low (0-3)		Iwate Intermed. (4-6)		Iwate High (7-12)		
Benchmark Cutoffs	**Non-BM Cases at BM-Centers (n=1104)**	**BM Cases at Non-BM-Centers (n=148)**	**Non-BM Cases at BM Centers (n=292)**	**BM Cases at Non-BM Centers (n=27)**	**Non-BM Cases at BM Centers (n=280)**	**BM Cases at Non-BM Centers (n=47)**	**Non-BM Cases at BM Centers (n=624)**	**BM Cases at Non-BM Centers (n=85)**	**Non-BM Cases at BM Centers (n=492)**	**BM Cases at Non-BM Centers (n=43)**	**Donor hepatectomy (n=1236)**
Perioperative Metrics											
Operation time (min)	180 [120-249]	230 [144-334]	300 [234-389]	491 [319-590]	141 [105-215]	168 [97-234]	185 [125-270]	248 [152-381]	250 [180-340]	349 [265-487]	390 [303-461]
Estimated blood loss (mL)	100 [50-200]	100 [30-200]	200 [100-458]	220 [100-475]	75 [30-200]	50 [15-100]	100 [50-200]	100 [50-253]	150 [70-400]	200 [66-450]	75 [50-100]
Intraoperative blood transfusion (units)	0 [0-0]	0 [0-0]	0 [0-0]	0 [0-0]	0 [0-0]	0 [0-0]	0 [0-0]	0 [0-0]	0 [0-0]	0 [0-0]	0 [0-0]
Open conversion (%)	40 (3.6)	2 (1.4)	20 (6.8)	4 (14.8)	3 (1.1)	2 (4.3)	26 (4.2)	3 (3.5)	31 (6.3)	1 (2.3)	2 (0.2)
Hospital stay (d)	5 [3-10]	4 [3-6]	6 [4-10]	8 [5-14]	3 [2-9]	3 [2-5]	5 [3-10]	5 [3-7]	5 [4-10]	6 [4-10]	4 [3-5]
Postoperative Morbidity at 90d											
Any complication (%)	156 (14.1)	13 (8.8)	89 (30.5)	8 (29.6)	24 (8.6)	5 (10.6)	105 (16.8)	8 (9.4)	116 (23.6)	8 (18.6)	107 (8.7)

Clavien–Dindo grade ≥3a (%)	≤2.2	7 (4.7)	≤5.9	6 (22.2)	10 (3.6)	4 (8.5)	38 (6.1)	5 (5.9)	47 (9.6)	4 (9.3)	16 (1.3)
CCI®	0 [0-0]	0 [0-0]	0 [0-21]	0 [0-21]	0 [0-0]	0 [0-0]	0 [0-0]	0 [0-0]	0 [0-0]	0 [0-0]	0 [0-0]
Relaparotomy rate (%)	61 (6.6)	1 (0.7)	12 (5.1)	1 (4.5)	18 (7.7)	1 (2.2)	34 (6.3)	1 (1.2)	21 (5.4)	0 (0.0)	7 (0.6)
Readmission rate (%)	58 (5.3)	6 (4.1)	32 (11.0)	0 (0.0)	16 (5.7)	4 (8.5)	38 (6.1)	2 (2.4)	36 (7.3)	0 (0.0)	9 (0.7)
Mortality (%)	8 (0.8)	0 (0.0)	5 (1.8)	1 (3.7)	0 (0.0)	0 (0.0)	8 (1.4)	1 (1.2)	5 (1.2)	0 (0.0)	0 (0)
Liver failure – ISGLS Grade B/C (%)	3 (0.3)	0 (0.0)	9 (3.1)	2 (7.4)	0 (0.0)	0 (0.0)	5 (0.8)	2 (2.4)	7 (1.4)	0 (0.0)	4 (0.3)
Bile leak – ISGLS Grade B/C (%)	25 (2.3)	3 (2.0)	29 (9.9)	5 (18.5)	4 (1.4)	0 (0.0)	19 (3.0)	5 (5.9)	31 (6.3)	3 (7.0)	24 (1.9)
Textbook outcome (%)	872 (80.1)	124 (83.8)	184 (63.0)	15 (55.6)	240 (85.7)	37 (78.7)	490 (78.5)	69 (81.2)	337 (68.5)	33 (76.7)	1195 (96.7)
Oncological Outcome											
R1 resection (%)	77 (7.0)	11 (7.5)	33 (11.3)	4 (14.8)	11 (3.9)	3 (6.5)	49 (7.9)	9 (10.6)	50 (10.2)	3 (7.0)	-

Abbreviations: BM, Benchmark; CCI®, Comprehensive Complication Index; ISGLS, International Study Group of Liver Surgery

Table 3: Multivariable Logistic Regression Analysis for Not Achieving Textbook Outcome.

Characteristic	OR	95% CI	p-value
Benchmark case	0.6	0.48, 0.71	<0.001
Malignant indication	5.0	3.89, 6.59	<0.001
Neoadjuvant chemotherapy	1.5	1.18, 1.87	<0.001
Major hepatectomy	1.6	1.29, 2.00	<0.001
Iwate score	1.1	1.03, 1.12	<0.001

Abbreviations: OR, Odds Ratio; CI, Confidence Interval.