



Original Investigation | Surgery

Nationwide Association of Surgical Performance of Minimally Invasive Esophagectomy With Patient Outcomes

Mirte H. M. Ketel, MSc; Bastiaan R. Klarenbeek, MD, PhD; Inger Abma, PhD; Eric H. J. Belgers, MD; Peter-Paul L. O. Coene, MD; Jan Willem T. Dekker, MD, PhD; Peter van Duijvendijk, MD, PhD; Marloes Emous, MD, PhD; Suzanne S. Gisbertz, MD, PhD; Jan Willem Haveman, MD, PhD; Joos Heisterkamp, MD, PhD; Grard A. P. Nieuwenhuijzen, MD, PhD; Jelle P. Ruurda, MD, PhD; Johanna W. van Sandick, MD, PhD; Pieter C. van der Sluis, MD, PhD; Marc J. van Det, MD, PhD; Stijn van Esser, MD, PhD; Simon Law, MD, PhD; Wobbe O. de Steur, MD, PhD; Meindert N. Sosef, MD, PhD; Bas Wijnhoven, MD, PhD; Gerjon Hannink, PhD; Camiel Rosman, MD, PhD; Frans van Workum, MD, PhD; for the MIE-CAT Netherlands Collaborative Group

Abstract

IMPORTANCE Suboptimal surgical performance is hypothesized to be associated with less favorable patient outcomes in minimally invasive esophagectomy (MIE). Establishing this association may lead to programs that promote better surgical performance of MIE and improve patient outcomes.

OBJECTIVE To investigate associations between surgical performance and postoperative outcomes after MIE.

DESIGN, SETTING, AND PARTICIPANTS In this nationwide cohort study of 15 Dutch hospitals that perform more than 20 MIEs per year, 7 masked expert MIE surgeons assessed surgical performance using videos and a previously developed and validated competency assessment tool (CAT). Each hospital submitted 2 representative videos of MIEs performed between November 4, 2021, and September 13, 2022. Patients registered in the Dutch Upper Gastrointestinal Cancer Audit between January 1, 2020, and December 31, 2021, were included to examine patient outcomes.

EXPOSURE Hospitals were divided into quartiles based on their MIE-CAT performance score. Outcomes were compared between highest (top 25%) and lowest (bottom 25%) performing quartiles. Transthoracic MIE with gastric tube reconstruction.

MAIN OUTCOME AND MEASURE The primary outcome was severe postoperative complications (Clavien-Dindo ≥ 3) within 30 days after surgery. Multilevel logistic regression, with clustering of patients within hospitals, was used to analyze associations between performance and outcomes.

RESULTS In total, 30 videos and 970 patients (mean [SD] age, 66.6 [9.1] years; 719 men [74.1%]) were included. The mean (SD) MIE-CAT score was 113.6 (5.5) in the highest performance quartile vs 94.1 (5.9) in the lowest. Severe postoperative complications occurred in 18.7% (41 of 219) of patients in the highest performance quartile vs 39.2% (40 of 102) in the lowest (risk ratio [RR], 0.50; 95% CI, 0.24-0.99). The highest vs the lowest performance quartile showed lower rates of conversions (1.8% vs 8.9%; RR, 0.21; 95% CI, 0.21-0.21), intraoperative complications (2.7% vs 7.8%; RR, 0.21; 95% CI, 0.04-0.94), and overall postoperative complications (46.1% vs 65.7%; RR, 0.54; 95% CI, 0.24-0.96). The R0 resection rate (96.8% vs 94.2%; RR, 1.03; 95% CI, 0.97-1.05) and lymph node yield (mean [SD], 38.9 [14.7] vs 26.2 [9.0]; RR, 3.20; 95% CI, 0.27-3.21) increased with oncologic-specific performance (eg, hiatus dissection, lymph node dissection). In addition, a high anastomotic phase score was associated with a lower anastomotic leakage rate (4.6% vs 17.7%; RR, 0.14; 95% CI, 0.06-0.31).

(continued)

Key Points

Question What is the association between surgical performance of minimally invasive esophagectomy (MIE) and patient outcomes?

Findings In this cohort study, surgical performance across all 15 Dutch hospitals performing MIE was associated with intraoperative and severe postoperative complications; conversion occurred less frequently in the highest compared with the lowest performance hospital quartile, and better anastomosis creation performance resulted in decreased anastomotic leakage.

Meaning These findings suggest that better patient outcomes on a national level may be achievable by improving surgical performance of MIE, likely with MIE competency assessment tool implementation.

+ Invited Commentary

+ Supplemental content

Author affiliations and article information are listed at the end of this article.

Open Access. This is an open access article distributed under the terms of the CC-BY-NC-ND License.

Abstract (continued)

CONCLUSIONS AND RELEVANCE These findings suggest that better surgical performance is associated with fewer perioperative complications for patients with esophageal cancer on a national level. If surgical performance of MIE can be improved with MIE-CAT implementation, substantially better patient outcomes may be achievable.

JAMA Network Open. 2024;7(4):e246556. doi:10.1001/jamanetworkopen.2024.6556

Introduction

There is a growing body of evidence showing substantial variation in surgical performance of various minimally invasive procedures.¹ Strong associations between surgical performance and patient outcomes of these various procedures have been reported.¹⁻⁸ Minimally invasive esophagectomy (MIE) is a technically challenging procedure, and it is hypothesized that variations in surgical performance (ie, technical conduct of the surgery) of MIE are associated with variations in patient outcomes.

Many factors substantiate the hypothesis that MIE performance is associated with outcome. For example, MIE is known to have a long learning curve of more than 100 cases,^{9,10} and an important volume-outcome association has been described.¹¹ More importantly, there is substantial national variation in patient outcome among hospitals^{12,13} for which we were unable to find explanations in differences in patient case mix, protocols, or other associated factors. If a substantial performance-outcome association for MIE exists, increasing surgical performance of MIE may be essential for improving patient outcomes.

Performance of minimally invasive procedures can be reliably measured by validated competency assessment tools (CATs).^{2,3,14,15} In a previous study, the MIE-CAT was developed and validated, and higher MIE-CAT scores were found to be associated with less operative time, less blood loss, and fewer intraoperative complications.¹⁶ However, it remains unknown whether surgical performance of MIE is associated with patient outcomes on a national level. Thus, the aim of this study was to investigate the association of surgical performance of transthoracic MIE and postoperative outcomes in the Netherlands.

Methods

Study Design

This nationwide cohort study used observational video analysis to assess the surgical performance of MIE in the Netherlands. All 15 Dutch hospitals that perform more than 20 robot-assisted or thoracoscopic MIEs per year were approached to participate. This study does not fall within the scope of the Dutch Medical Research Involving Human Subjects Act. We obtained an official exemption from the local medical ethics committee CMO Arnhem-Nijmegen and participating hospitals. This study was reported according to the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) guideline.

Each participating hospital received a survey regarding the surgical team's experience and primary MIE technique and was invited to submit their patient outcomes from January 1, 2020, through December 31, 2021, and 2 full-length MIE videos filmed between November 4, 2021, and September 13, 2022, representative of their hospital's primary MIE technique in 2020-2021. This procedure is in line with previous studies,^{2,3,14,15} which reported meaningful results using this approach. Although we recognized the additional value of linking performance of individual videos to outcome of corresponding patients before starting this study, this workload was deemed unfeasible.

Surgical Procedures and Videos

The recorded esophagectomies were complete MIEs performed transthoracically using a conventional minimally invasive (ie, thoracoscopic) or robot-assisted approach, with gastric tube reconstruction and intrathoracic (Ivor-Lewis¹⁷) or cervical (McKeown¹⁸) anastomosis. Other approaches were excluded (eg, Orringer esophagectomy¹⁹ or minimally invasive transcervical esophagectomy²⁰). Participating hospitals were asked to record full-length procedures, including outside the patient's body if applicable (eg, the handsewn cervical anastomosis, extracorporeal gastric tube creation). Videos were included after written informed patient consent and stripped of any patient or surgeon identifiers.

Study Population

Data on patients who underwent MIE, as registered in the mandatory national Dutch Upper Gastrointestinal Cancer Audit (DUCA) registry,²¹ were provided by each hospital. Patients were included if they underwent complete minimally invasive transthoracic MIE with gastric tube reconstruction for curative treatment between January 1, 2020, and December 31, 2021, corresponding to the most recently available data at the time of the study. Clinically relevant parameters before, during, and after the operation, such as patient characteristics, surgery date, the occurrence and severity of complications, and lymph node yield, were extracted from the DUCA registry.

Reviewers and Surgical Performance Assessment

The previously developed and validated procedure-specific MIE-CAT¹⁶ consists of 8 phases, with add-ons for various intrathoracic and cervical anastomotic techniques (eMethods 1 in [Supplement 1](#)). Each phase contains 4 quality components (exposure, execution, adverse events, and end-product quality) scored on a 1 to 4 Likert scale, with higher scores indicating better performance. A total MIE-CAT performance score is calculated as the sum of all phases, with a range of 32 to 128 points. Since national and international consensus regarding radicality of the lymphadenectomy is lacking,²²⁻²⁴ the standard extent of lymphadenectomy can be tailored in the MIE-CAT. In the present study, the standard was derived from a previous Delphi study²⁵ (eMethods 1 in [Supplement 1](#)).

Seven masked and independent MIE surgeons assessed surgical performance. Three expert consultant surgeons, 2 (M.J.v.D. and S.v.E.) from the Netherlands and 1 international (S.L.) from Hong Kong, assessed the full-length MIE videos. These surgeons had a minimum of 5 years' experience with MIE, had performed at least 120 MIE procedures,⁹ and were each experienced in 1 of 3 primary anastomotic techniques: intrathoracic end-to-side, intrathoracic side-to-side, and cervical hand-sewn end-to-end. Considering that the anastomosis is a clinically relevant phase,²⁶ 4 additional Dutch expert surgeons (B.R.K., W.O.d.S., M.N.S., and B.W.) with a minimum of 5 years' experience and extensive expertise in their preferred anastomotic technique conducted the anastomotic phase assessment of their daily practice-preferred anastomosis. All reviewers received a 1-hour online workshop prior to the assessments during which assessing performance with the MIE-CAT was explained and practice was conducted with video clips.

Study Parameters and Outcome Measures

All videos and DUCA data were pseudonymized for data analysis. Each hospital's surgical performance was calculated as the mean total MIE-CAT score of the 2 submitted videos (eMethods 1 in [Supplement 1](#)). The primary outcome was the occurrence of any severe postoperative complication with a Clavien-Dindo classification of 3 or higher within 30 days after surgery. Secondary outcomes included intraoperative complication (during the procedure), any postoperative complication, anastomotic leakage, pulmonary complication, RO resection, readmission, reintervention, 30-day mortality, textbook outcome,²⁷ intensive care unit (ICU) length of stay, and lymph node yield.²⁷

Statistical Analysis

Hospitals were divided into quartiles based on their performance score, including high (quartile 4, top 25%), medium (quartiles 2 and 3, middle 50%), and low (quartile 1, bottom 25%). Multivariable, multilevel logistic regression analysis was used to assess associations among surgical performance, MIE-CAT (component) score, and clinical outcomes between the highest and lowest performance quartiles. Associations to be analyzed were chosen beforehand and included the association between anastomotic phase performance (MIE-CAT phase 8) and both anastomotic leakage and textbook outcome, between end-product quality of lymph node dissection (MIE-CAT phases 3 and 7) and lymph node yield, between end-product quality of hiatus dissection (phase 4) and thoracic esophageal mobilization (phase 6) and RO resection rate, and between both total adverse events and total end-product quality and ICU length of stay. Associations are presented as risk ratios (RRs) converted from odds ratios²⁸ and as absolute risk differences with 95% CIs. In the multilevel models, hospital was included as a random intercept, and 11 patient- and performance score-related confounders were selected, including sex, age, body mass index, American Society of Anesthesiologists (ASA) classification, clinical T and N classification, tumor location, anastomosis type (ie, intrathoracic, cervical), neoadjuvant therapy, and abdominal and thoracic case difficulty (eMethods 2 in Supplement 1). Categorical confounders were clustered if fewer than 10 events per category were observed (eg, ASA 1 and 2, ASA 3 and 4). Surgical performance was analyzed with regard to MIE experience in years and volume (procedures over 2 years per hospital) using Pearson correlation coefficients. Intraclass correlation coefficients (ICCs), using a 2-way random effect for interrater reliability and a 2-way mixed effect for test-retest reliability between external and Dutch experts, were calculated for external validation of the MIE-CAT. The data analysis was performed using SPSS Statistics for Windows, version 27.0 (IBM Corporation) and package lme4 in R, version 4.3.0 (R Foundation for Statistical Computing).

Results

Hospitals and Surgeons

All 15 Dutch hospitals that perform MIE participated in this study (eTable 1 in Supplement 1). Mean surgical experience with MIE was 10 years (IQR, 4-19 years). The mean number of attending esophageal surgeons was 3 per hospital (range, 2-5 per hospital), and the majority were visible in the submitted videos (eTable 2 in Supplement 1).

Patients and Outcomes

A total of 970 patients who underwent transthoracic MIE between 2020 and 2021 (mean [SD] age, 66.6 [9.1] years; 719 men [74.1%] and 251 women [25.9%]) were included. Patient characteristics are reported in eTable 3 in Supplement 1. Conversion, intraoperative complications, and severe postoperative complications occurred in 34 patients (3.5%), 42 patients (4.3%), and 121 patients (12.5%), respectively.

Surgical Video Assessments

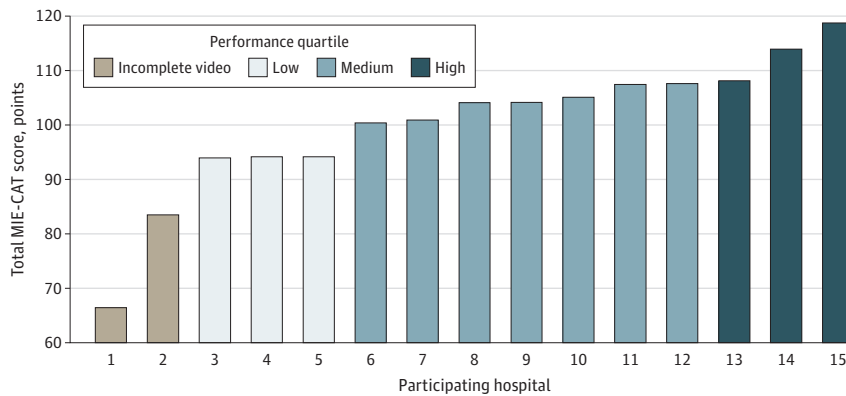
A total of 30 videos of operations from the 15 hospitals were included. Overall, 25 videos were complete and 5 from 3 hospitals had missing components: 2 videos were missing phase 5 (creation of the gastric tube), and 3 videos were missing both phase 5 and phase 8 (creation of the anastomosis). Together, a total MIE-CAT score was calculated for 13 hospitals. Two hospitals with missing components on both videos were excluded for the total MIE-CAT score but included for the MIE-CAT component scores that were available. The total MIE-CAT scores ranged from 93.9 to 118.8 points (mean [SD], 104.1 [7.5] points) (Figure 1). The mean (SD) MIE-CAT score was 113.6 (5.5) points in the highest performance group, 103.3 (6.3) in the medium performance group, and 94.1 (5.9) in the lowest performance group table 1 (difference between lowest and highest performance quartiles, 19.5; 95% CI, 12.2-26.9).

Surgical Performance of MIE and Clinical Outcomes

In the highest performance quartile, more favorable clinical outcomes were seen compared with the lowest performance quartile (Table 1; Figure 2). Severe postoperative complications occurred in 18.7% (41 of 219) of the patients in the highest compared with 39.2% (40 of 102) in the lowest performance quartile (Table 1). In multilevel analysis, a higher mean MIE-CAT score was significantly associated with fewer severe postoperative complications (RR, 0.50; 95% CI, 0.24-0.99), with a 20.0% absolute risk in severe postoperative complications in the highest vs 35.8% in the lowest performance quartile (absolute risk difference, 15.8%; 95% CI, 3.5%-28.0%). Performance of the highest quartile was also associated with a 1.8% absolute risk of conversion and 2.4% risk of intraoperative complications vs 7.6% (absolute risk difference, 5.8%; 95% CI, 0.3%-15.1%) and 10.2% (absolute risk difference, 7.8%; 95% CI, 1.3%-21.7%), in the lowest performance quartile, respectively (Figure 2). Overall, within the highest vs the lowest performance quartile, lower rates of conversions (1.8% vs 8.9%; RR, 0.21; 95% CI, 0.21-0.21), intraoperative complications (2.7% vs 7.8%; RR, 0.21; 95% CI, 0.04-0.94), and overall postoperative complications (46.1% vs 65.7%; RR, 0.54; 95% CI, 0.24-0.96) were seen (Table 1).

Other associations were in line with the hypothesized direction of effect (ie, higher MIE-CAT score associated with better patient outcomes) but were not significant (eg, reintervention RR, 0.55;

Figure 1. Distribution of Minimally Invasive Esophagectomy Competency Assessment Tool (MIE-CAT) Performances Scores



Total MIE-CAT scores range from 32 to 128, with higher scores indicating better performance.

Table 1. Patient Outcomes by Performance Quartile and the Associations Between Lowest and Highest Quartiles^a

Patient outcome	Performance quartile, ^b No. (%)			Low vs high, RR (95% CI) ^c
	Low (n = 102)	Medium (n = 617)	High (n = 219)	
Severe postoperative complication (Clavien-Dindo ≥3)	40 (39.2)	174 (28.2)	41 (18.7)	0.50 (0.24-0.99)
Postoperative complication	67 (65.7)	378 (61.3)	101 (46.1)	0.54 (0.24-0.96)
Anastomotic leakage	18 (17.7)	91 (14.8)	10 (4.6)	0.36 (0.10-1.20)
Pulmonary complication	28 (27.5)	185 (30.0)	47 (21.5)	0.66 (0.38-1.09)
Intraoperative complication	8 (7.8)	28 (4.5)	6 (2.7)	0.21 (0.04-0.94)
Conversion	9 (8.9) ^d	20 (3.2)	4 (1.8)	0.21 (0.21-0.21)
Reintervention	27 (26.5)	168 (27.4) ^d	35 (16.0)	0.55 (0.27-1.03)
ICU stay, median (IQR), d	3.4 (1.7-4.3)	3.4 (2.2-4.1)	2.2 (1.9-2.3)	0.22 (0.03-1.55)
Readmission	22 (22) ^d	104 (17.5) ^d	32 (14.7) ^d	0.97 (0.50-1.75)
Textbook outcome	51 (50.0)	305 (49.4)	144 (65.8)	1.45 (0.95-9.98)
30-Day mortality	3 (2.9)	16 (2.6)	8 (3.7)	1.44 (0.26-6.82)
R0 resection	101 (95.1) ^d	589 (95.9) ^d	182 (94.7) ^d	1.01 (0.85-1.06)
Lymph node yield, median (IQR)	30.1 (28.1-30.7)	28.7 (27.2-38.9)	28.8 (26.2-30.1)	0.20 (0.00-3.21)

Abbreviations: ICU, intensive care unit; RR, risk ratio.

^a Crude patient outcomes from the Dutch Upper Gastrointestinal Cancer Audit. In total, 938 of the 970 patients were used for the total minimally invasive esophagectomy comprehensive assessment tool analysis, as patients from the hospitals with incomplete videos were excluded.

^b Low, quartile 1; medium, quartiles 2 and 3; high, quartile 4.

^c Relative risk associations between the lowest and highest performance quartiles derived from a multilevel model.

^d These variables have some missing values in the Dutch Upper Gastrointestinal Cancer Audit.

95% CI, 0.27-1.03). No association between total MIE-CAT score and 30-day mortality was found (RR, 1.44; 95% CI, 0.26-6.82) (Table 1).

Regarding oncologic outcomes, R0 resection and lymph node yield were comparable between the highest and lowest performance quartiles based on total MIE-CAT scores (Table 1). Interestingly, **Table 2** shows that the R0 resection rate was 96.8% in the quartile with the highest end-product quality scores of hiatus dissection (phase 4) and mobilization of the thoracic esophagus (phase 6) vs 94.2% in the lowest quartile (RR, 1.03; 95% CI, 0.97-1.05) (eTable 4 in [Supplement 1](#)). Similarly, in the highest performance quartile for the end-product quality scores of both abdominal (phase 3) and thoracic (phase 7) lymph node dissection, the mean (SD) lymph node yield was 38.9 (14.7) compared with 26.2 (9.0) in the lowest quartile (RR, 3.19 [95% CI, 0.01-3.21] and 3.21, [95% CI, 3.21-3.21], respectively).

Higher performance of the creation of the anastomosis (phase 8) was also associated with less anastomotic leakage (highest vs lowest performance quartile rate, 4.6% vs 17.7%; RR, 0.14; 95% CI, 0.06-0.31) (Table 1; eTable 5 in [Supplement 1](#)). Again, nearly all procedure-specific component performance scores and outcomes were in line with our hypothesis that improved performance is associated with improved outcome (eTable 5 in [Supplement 1](#)). In addition, some associations were found for both total adverse events score and total end-product quality score with ICU length of stay (RR, 0.24 [95% CI, 0.06-1.00] and 0.17 [95% CI, 0.03-1.00], respectively).

Additional Analyses

A moderate positive correlation between surgical performance and experience with MIE in years was found ($r = 0.61$, 95% CI, 0.30-0.81) (eFigure 1 in [Supplement 1](#)). Importantly, the association between surgical performance of MIE and severe postoperative complications was still found when excluding

Figure 2. Absolute Risks in Obtaining Clinical Outcomes Based on Multilevel Analysis of Surgical Performance

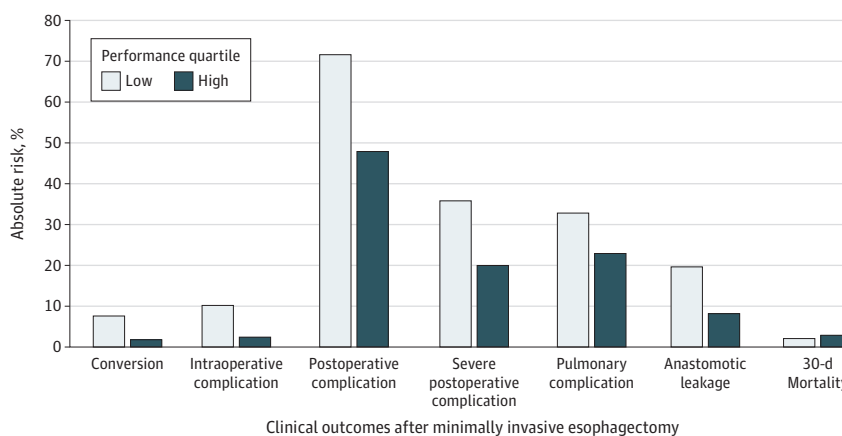


Table 2. Oncologic Outcomes per Performance Quartile of MIE-CAT Component Scores

MIE-CAT component score by oncologic outcome	Performance quartile ^a		
	Low	Medium	High
R0 resection rate, No. (%)			
Hiatus dissection (EPQ phase 4)	72 (96.0)	537 (95.6)	288 (95.1)
Thoracic esophagus mobilization (EPQ phase 6)	122 (93.1)	458 (95.8)	317 (95.8)
Hiatus dissection and thoracic esophagus mobilization (EPQ phases 4 and 6)	195 (94.2)	434 (95.2)	268 (96.8)
Lymph node yield, mean (SD)			
Abdominal lymph node dissection (EPQ phase 3)	26.0 (8.9)	27.9 (9.7)	38.1 (14.6)
Thoracic lymph node dissection (EPQ phase 7)	27.6 (8.8)	27.2 (9.7)	38.1 (14.6)
Complete lymph node dissection (EPQ phases 3 and 7)	26.2 (9.0)	28.1 (9.9)	38.9 (14.7)

Abbreviations: EPQ, end-product quality; MIE-CAT, minimally invasive esophagectomy competency assessment tool.

^a Low, quartile 1; medium, quartiles 2 and 3; high, quartile 4.

2 centers that might not have passed the learning curve of 119 cases⁹ (RR, 0.96; 95% CI, 0.93-0.99). There was a weak positive correlation between surgical performance and MIE volume (procedures over 2 years per hospital: $r = 0.45$; 95% CI, 0.07-0.71) (eFigure 2 in Supplement 1). Surprisingly, mean hospital volume was highest in the medium performance quartiles instead of in the highest performance quartile (Table 3). No differences were found between laparoscopic and robot-assisted MIE (eFigure 3 in Supplement 1).

External Validation of the MIE-CAT

Comparable to the validation study of the MIE-CAT,¹⁶ interrater reliability between reviewers was acceptable (phase ICC, 0.69 [95% CI, 0.65-0.94]; quality component ICC, 0.78 [95% CI, 0.57-0.97]; total MIE-CAT score ICC, 0.69 [95% CI, 0.20-0.97]), and test-retest reliability was good (ICC, 0.86; 95% CI, 0.58-0.95). Moreover, scores of the 2 videos from each hospital were similar ($r = 0.77$; 95% CI, 0.42-0.92).

Performance assessments of the international expert (mean total MIE-CAT [SD], 110.8 [10.7] points) and the 2 Dutch experts (mean total MIE-CAT [SD], 106.3 [6.6] points) showed comparable scores (mean [SD] difference, 4.5 [7.8] points) (eTable 6 in Supplement 1) and acceptable interrater reliability (ICC, 0.77; 95% CI, 0.42-0.91). However, additional analysis showed that the international expert scores correlated more with lymph node yield compared with the Dutch expert scores (5 of 7 positive correlations vs 3 of 7) (eTable 4 in Supplement 1).

Discussion

This nationwide cohort study found substantial variation in surgical performance of MIE among Dutch hospitals and a strong association between surgical performance and clinical outcome. Increased surgical performance was associated with a substantial risk reduction of intraoperative and severe (Clavien-Dindo ≥ 3) postoperative complications and a lower conversion rate. A higher anastomotic phase performance score was associated with decreased anastomotic leakage. RO resection rate and lymph node yield were higher when relevant procedural phases of the MIE-CAT

Table 3. Surgical Performance and Experience by Quartile

	Performance quartile, ^a mean (SD)			Mean difference between low and high (95% CI)
	Low (n = 3)	Medium (n = 7)	High (n = 3)	
Surgical performance (total MIE-CAT score)	94.1 (5.9)	103.3 (6.3)	113.6 (5.5)	19.5 (12.2 to 26.9)
Phase 1: greater curvature mobilization	11.3 (1.4)	12.8 (1.3)	14.3 (0.8)	3.0 (1.5 to 4.4)
Phase 2: lesser curvature mobilization	11.1 (2.0)	13.1 (1.3)	14.9 (0.9)	3.8 (1.8 to 5.8)
Phase 3: abdominal lymph node dissection	11.0 (1.5)	13.1 (1.3)	14.1 (1.6)	3.1 (1.2 to 5.1)
Phase 4: hiatus dissection	11.7 (2.1)	13.3 (1.0)	14.5 (1.0)	2.8 (0.6 to 4.9)
Phase 5: gastric tube creation	12.6 (1.3)	12.6 (3.8)	14.4 (1.2)	1.8 (0.2 to 3.3)
Phase 6: thoracic esophagus mobilization	12.7 (1.4)	13.2 (1.2)	13.5 (1.3)	0.8 (-1.0 to 2.5)
Phase 7: thoracic lymph node dissection	13.0 (1.3)	13.4 (1.1)	13.9 (0.8)	0.9 (-0.5 to 2.3)
Phase 8: anastomosis creation	10.6 (4.1)	11.9 (1.1)	14.1 (1.1)	3.5 (-0.3 to 7.3)
Exposure	22.4 (2.0)	24.5 (1.8)	27.5 (1.8)	5.1 (2.7 to 7.5)
Execution	23.0 (1.4)	25.3 (2.0)	27.8 (1.8)	4.8 (2.7 to 6.8)
Adverse events	25.1 (2.2)	27.0 (2.1)	28.9 (1.1)	3.8 (1.6 to 6.0)
End-product quality	24.3 (1.4)	27.6 (1.6)	29.5 (1.4)	5.2 (3.4 to 7.0)
Experience				
Experience with MIE, mean (SD), y	6.7 (2.3)	10.6 (2.2)	12.0 (6.6)	5.3 (-5.8 to 16.5)
MIE volume, No. (%)	34.3 (15.9)	88.9 (43.3)	73.0 (25.2)	38.7 (-9.1 to 86.5)

Abbreviations: MIE-CAT, minimally invasive esophagectomy competency assessment tool.

^a Low, quartile 1; medium, quartiles 2 and 3; high, quartile 4.

were performed better. The study findings suggest that surgical performance is an important target for programs to further improve patient outcomes after MIE.

Even though the association between surgical performance and clinical outcomes of MIE has been widely hypothesized,¹¹ this study is the first to report results supporting this hypothesis. Our findings are in line with previous surgical performance studies of various other procedures,^{2,3,5,6,15} but the first to be conducted nationwide. In addition, our study findings correlate well with previously established effect sizes of surgical learning curve studies.^{9,11} Even after excluding hospitals that may not have passed the learning curve,²⁹ performance and severe postoperative complications remained associated, indicating that the associations found in this study are not just a learning curve issue. Interestingly, hospital volume was the highest in the medium performance quartile hospitals, which suggests that our findings may not be solely explained by the volume-outcome association.¹¹

Oncologic outcomes were associated with end-product quality of relevant phases (ie, hiatus dissection and thoracic esophagus mobilization for RO resection, abdominal and thoracic lymph node dissection for lymph node yield). Interestingly, the scores of the international expert from Hong Kong correlated more with lymph node yield compared with the Dutch experts, which may indicate that there is merit to the more radical school of thought among eastern MIE surgeons³⁰ and that surgeons worldwide may benefit from learning from one another, although further study of this hypothesis is needed.

Our findings have important clinical implications for practices inside and outside of Dutch hospitals. First, establishing that surgical performance is associated with clinical outcome comes with the opportunity and obligation to improve performance. Assessing surgical performance and providing direct feedback has been shown to be a powerful and effective tool in other surgical procedures.³¹⁻³⁴ Therefore, we have initiated a surgical performance improvement program in which videos are assessed using the MIE-CAT and discussed within groups of surgeons from different hospitals. We will refine and evaluate this program to investigate whether it can lead to improved surgical performance and, ultimately, patient outcomes. If successful, we aim to expand this program. Second, standardized video collection, preferably by a trusted national outcome registry (eg, the DUCA registry), might unlock a great potential for research and quality improvement. Moreover, video databases may have great value in artificial intelligence training,^{35,36} especially when combining CAT scores and patient outcomes. Finally, a CAT may be updated as additional data become available, such as regarding the association of surgical details (eg, gastric tube width) with outcomes or with new consensus statements (eg, lymphadenectomy standards). Additionally, assessment efficiency may be improved by scoring the most important procedural phases and the use of artificial intelligence. Further investigation is needed on whether these innovations might provide comparable results.

Strengths and Limitations

Major strengths of this study are the nationwide design, high-quality national registry data, and a thoroughly validated CAT for performance assessment and video review by MIE experts, including an international expert. The open and collaborative culture of Dutch esophageal surgeons has been of paramount importance to the successful conduct of this study. Furthermore, the study was used for external validation of the MIE-CAT.

This study also has several limitations. The most important limitation was the 2 self-selected representative videos of the hospital's surgical performance during the 2-year study period. Ideally, videos from all patients would have been analyzed in relation to the individual patient's outcome and reported per operating surgeon. This heavy workload was deemed unfeasible with limited time and resources. However, the representation of videos for a hospital may have influenced the study results. Although the limited number of surgeons per hospital and the fact that MIE is often performed by 2 consultant surgeons strengthen our belief that the videos are representative of a hospital's technical performance, we acknowledge this uncertainty. Furthermore, the hospitals' self-selected videos may have introduced selection bias (eg, selecting videos of surgeries that went

better than average), potentially resulting in less variation between hospitals and bias toward the null hypothesis. In addition, DUCA data did not include overall care parameters and were not available at the surgeon level. Therefore, we assumed comparable quality of care and performance between surgeons of a hospital. The fact that we found important correlations despite these design limitations may, in our view, be an outcome of the strong association between surgical performance and patient outcome established for other procedures.^{2,3,5,15} Following this line of reasoning, the strong associations between performance and outcomes may be an underestimation of the actual effect, which warrants investigation in future studies. Another important limitation is the potential variation in reviewer opinion on good performance introducing bias, particularly regarding lymphadenectomy. Without consensus on the extent of lymphadenectomy,¹⁶ reviewers were asked to score all visible lymph node stations in the videos, and only consensus-based stations from the Delphi study²⁵ counted for the MIE-CAT score. Although acceptable reliability between reviewers (ICC >0.7) was established, consensus on the extent of lymphadenectomy may further improve reliability of the MIE-CAT.

Conclusions

In this study, better surgical performance was associated with better patient outcomes on a national level in the Netherlands. Significantly better patient outcomes may be achievable if surgical performance of MIE can be improved.

ARTICLE INFORMATION

Accepted for Publication: January 31, 2024.

Published: April 19, 2024. doi:10.1001/jamanetworkopen.2024.6556

Open Access: This is an open access article distributed under the terms of the [CC-BY-NC-ND License](#). © 2024 Ketel MHM et al. *JAMA Network Open*.

Corresponding Author: Frans van Workum, MD, PhD, Department of Surgery, Radboud University Medical Center, PO Box 9101, 6500 HB, Nijmegen, the Netherlands (frans.vanworkum@radboudumc.nl).

Author Affiliations: Department of Surgery, Radboud University Medical Center, Nijmegen, the Netherlands (Ketel, Klarenbeek, Rosman, van Workum); IQ Healthcare, Radboud University Medical Center, Nijmegen, the Netherlands (Abma); Department of Surgery, Zuyderland, Heerlen, the Netherlands (Belgers, Sosef); Department of Surgery, Maastad Hospital, Rotterdam, the Netherlands (Coene); Department of Surgery, Reinier de Graaf Groep, Delft, the Netherlands (Dekker, van Esser); Department of Surgery, Gelre Ziekenhuizen, Apeldoorn, the Netherlands (van Duijvendijk); Department of Surgery, Medical Center Leeuwarden, Leeuwarden, the Netherlands (Emous); Department of Surgery, Amsterdam University Medical Center, Location University of Amsterdam, Amsterdam, the Netherlands (Gisbertz); Cancer Center Amsterdam, Cancer Treatment and Quality of Life, Amsterdam, the Netherlands (Gisbertz); Department of Surgery, University Medical Centre Groningen, University of Groningen, Groningen, the Netherlands (Haveman); Department of Surgery, Elisabeth Twee-Steden Hospital, Tilburg, the Netherlands (Heisterkamp); Department of Surgery, Catharina Hospital, Eindhoven, the Netherlands (Nieuwenhuijzen); Department of Surgery, University Medical Center Utrecht, Utrecht University, Utrecht, the Netherlands (Ruurda); Department of Surgery, Antoni van Leeuwenhoek Ziekenhuis, Amsterdam, the Netherlands (van Sandick, Wijnhoven); Department of Surgery, Erasmus MC University Medical Center Rotterdam, Rotterdam, the Netherlands (van der Sluis); Department of Surgery, Hospital Group Twente (ZGT), Almelo, the Netherlands (van Det); Department of Surgery, Queen Mary Hospital, School of Clinical Medicine, The University of Hong Kong, China (Law); Department of Surgery, Leiden University Medical Center, Leiden, the Netherlands (de Steur); Department of Medical Imaging, Radboud University Medical Center, Nijmegen, the Netherlands (Hannink); Department of Surgery, Canisius-Wilhelmina Hospital, Nijmegen, the Netherlands (van Workum).

Author Contributions: Ms Ketel and Dr van Workum had full access to all of the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis. Prof Rosman and Dr van Workum shared last authorship and contributed equally.

Concept and design: Ketel, Klarenbeek, Belgers, Coene, Dekker, Gisbertz, Haveman, Ruurda, van Esser, Rosman, van Workum.

Acquisition, analysis, or interpretation of data: Ketel, Klarenbeek, Abma, van Duijvendijk, Emous, Gisbertz, Haveman, Heisterkamp, Nieuwenhuijzen, van Sandick, van der Sluis, van Det, Law, de Steur, Sosef, Wijnhoven, Hannink, Rosman, van Workum.

Drafting of the manuscript: Ketel, Klarenbeek, Belgers, van der Sluis, Hannink, Rosman, van Workum.

Critical review of the manuscript for important intellectual content: Ketel, Klarenbeek, Abma, Coene, Dekker, van Duijvendijk, Emous, Gisbertz, Haveman, Heisterkamp, Nieuwenhuijzen, Ruurda, van Sandick, van der Sluis, van Det, van Esser, Law, de Steur, Sosef, Wijnhoven, Hannink, Rosman, van Workum.

Statistical analysis: Ketel, Abma, Hannink.

Obtained funding: Rosman, van Workum.

Administrative, technical, or material support: Ketel, Klarenbeek, Heisterkamp, van Sandick, van der Sluis, Wijnhoven, van Workum.

Supervision: Klarenbeek, Abma, Coene, Dekker, Emous, Gisbertz, Haveman, Heisterkamp, Ruurda, van der Sluis, van Det, Sosef, Hannink, Rosman, van Workum.

Conflict of Interest Disclosures: Dr Klarenbeek reported receiving grants from Medtronic, Stichting Bergh in het Zadel, ZonMw, and BENEFIT outside the submitted work. Dr van Det reported receiving proctor and consulting fees from Intuitive Surgical outside the submitted work. Dr Wijnhoven reported receiving grants and speaker's fees paid to his institution from Bristol-Myers Squibb outside the submitted work. Dr van Workum reported receiving grants and other support from Medtronic outside the submitted work. No other disclosures were reported.

Funding/Support: This study was funded by grants from Ethicon Endo-Surgery and Johnson & Johnson Medical Limited.

Role of the Funder/Sponsor: The study was performed independently, and Ethicon had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit to manuscript for publication.

Group Information: The MIE-CAT Netherlands Collaborative Group members are listed in [Supplement 2](#).

Data Sharing Statement: See [Supplement 3](#).

Additional Contributions: The authors thank all the people and the participating hospitals in the MIE-CAT Netherlands Collaborative Group (eAppendix in [Supplement 1](#)). The enlisted collaborators were all part of this study and were given the opportunity to review and edit the manuscript.

REFERENCES

1. Balvardi S, Kammili A, Hanson M, et al. The association between video-based assessment of intraoperative technical performance and patient outcomes: a systematic review. *Surg Endosc*. 2022;36(11):7938-7948. doi:10.1007/s00464-022-09296-6
2. Birkmeyer JD, Finks JF, O'Reilly A, et al; Michigan Bariatric Surgery Collaborative. Surgical skill and complication rates after bariatric surgery. *N Engl J Med*. 2013;369(15):1434-1442. doi:10.1056/NEJMsa1300625
3. Stulberg JJ, Huang R, Kreutzer L, et al. Association between surgeon technical skills and patient outcomes. *JAMA Surg*. 2020;155(10):960-968. doi:10.1001/jamasurg.2020.3007
4. Fecso AB, Bhatti JA, Stotland PK, Quereshy FA, Grantcharov TP. Technical performance as a predictor of clinical outcomes in laparoscopic gastric cancer surgery. *Ann Surg*. 2019;270(1):115-120. doi:10.1097/SLA.0000000000002741
5. Curtis NJ, Foster JD, Miskovic D, et al. Association of surgical skill assessment with clinical outcomes in cancer surgery. *JAMA Surg*. 2020;155(7):590-598. doi:10.1001/jamasurg.2020.1004
6. Chhabra KR, Thumma JR, Varban OA, Dimick JB. Associations between video evaluations of surgical technique and outcomes of laparoscopic sleeve gastrectomy. *JAMA Surg*. 2021;156(2):e205532. doi:10.1001/jamasurg.2020.5532
7. Mackenzie H, Ni M, Miskovic D, et al. Clinical validity of consultant technical skills assessment in the English National Training Programme for Laparoscopic Colorectal Surgery. *Br J Surg*. 2015;102(8):991-997. doi:10.1002/bjs.9828
8. Goldenberg MG, Goldenberg L, Grantcharov TP. Surgeon performance predicts early continence after robot-assisted radical prostatectomy. *J Endourol*. 2017;31(9):858-863. doi:10.1089/end.2017.0284
9. van Workum F, Stenstra MHBC, Berkelmans GHK, et al. Learning curve and associated morbidity of minimally invasive esophagectomy: a retrospective multicenter study. *Ann Surg*. 2019;269(1):88-94. doi:10.1097/SLA.0000000000002469
10. Claassen L, van Workum F, Rosman C. Learning curve and postoperative outcomes of minimally invasive esophagectomy. *J Thorac Dis*. 2019;11(suppl 5):S777-S785. doi:10.21037/jtd.2018.12.54

11. Claassen L, Hannink G, van Workum F, Rosman C. Response to the comment on "Learning Curves of Ivor Lewis Totally Minimally Invasive Esophagectomy by Hospital and Surgeon Characteristics a Retrospective Multi-National Cohort Study". *Ann Surg*. 2021;274(6):e930. doi:[10.1097/SLA.00000000000005147](https://doi.org/10.1097/SLA.00000000000005147)
12. Voeten DM, van der Werf LR, Wijnhoven BPL, van Hillegersberg R, van Berge Henegouwen MI; Dutch Upper GI Cancer Audit Group. Failure to cure in patients undergoing surgery for esophageal carcinoma: hospital of surgery influences prospects for cure: a nation-wide cohort study. *Ann Surg*. 2020;272(5):744-750. doi:[10.1097/SLA.0000000000004178](https://doi.org/10.1097/SLA.0000000000004178)
13. Nederlandse Kankerregistratie (NKR) IKNL, DUCA. Slokdarm- en maagkanker in Nederland. Integraal Kankercentrum Nederland. September 2021. Accessed April 1, 2022. https://iknl.nl/getmedia/196f17c1-3c86-41b8-ad9d-721ab0ba81d8/slokdarm_maag_def.pdf
14. Scally CP, Varban OA, Carlin AM, Birkmeyer JD, Dimick JB; Michigan Bariatric Surgery Collaborative. Video ratings of surgical skill and late outcomes of bariatric surgery. *JAMA Surg*. 2016;151(6):e160428. doi:[10.1001/jamasurg.2016.0428](https://doi.org/10.1001/jamasurg.2016.0428)
15. Varban OA, Thumma JR, Finks JF, Carlin AM, Ghaferi AA, Dimick JB. Evaluating the effect of surgical skill on outcomes for laparoscopic sleeve gastrectomy: a video-based study. *Ann Surg*. 2021;273(4):766-771. doi:[10.1097/SLA.0000000000003385](https://doi.org/10.1097/SLA.0000000000003385)
16. Ketel MH, Klarenbeek BR, Eddahchouri Y, et al A video-based procedure-specific competency assessment tool for minimally invasive esophagectomy. *JAMA Surg*. Published online December 27, 2023. doi:[10.1001/jamasurg.2023.6522](https://doi.org/10.1001/jamasurg.2023.6522)
17. Lewis I. The surgical treatment of carcinoma of the oesophagus; with special reference to a new operation for growths of the middle third. *Br J Surg*. 1946;34(133):18-31. doi:[10.1002/bjs.18003413304](https://doi.org/10.1002/bjs.18003413304)
18. McKeown KC. Total three-stage oesophagectomy for cancer of the oesophagus. *Br J Surg*. 1976;63(4):259-262. doi:[10.1002/bjs.1800630403](https://doi.org/10.1002/bjs.1800630403)
19. Orringer MB, Sloan H. Esophagectomy without thoracotomy. *J Thorac Cardiovasc Surg*. 1978;76(5):643-654. doi:[10.1016/S0022-5223\(19\)41012-X](https://doi.org/10.1016/S0022-5223(19)41012-X)
20. Klarenbeek BR, Fujiwara H, Scholte M, Rovers M, Shiozaki A, Rosman C. Introduction of minimally invasive transcervical oesophagectomy (MICE) according to the IDEAL framework. *Br J Surg*. 2023;110(9):1096-1099. doi:[10.1093/bjs/znad079](https://doi.org/10.1093/bjs/znad079)
21. Busweiler LA, Wijnhoven BP, van Berge Henegouwen MI, et al; Dutch Upper Gastrointestinal Cancer Audit (DUCA) Group. Early outcomes from the Dutch upper gastrointestinal cancer audit. *Br J Surg*. 2016;103(13):1855-1863. doi:[10.1002/bjs.10303](https://doi.org/10.1002/bjs.10303)
22. Hagens ERC, van Berge Henegouwen MI, Cuesta MA, Gisbertz SS. The extent of lymphadenectomy in esophageal resection for cancer should be standardized. *J Thorac Dis*. 2017;9(suppl 8):S713-S723. doi:[10.21037/jtd.2017.07.42](https://doi.org/10.21037/jtd.2017.07.42)
23. Hagens ERC, van Berge Henegouwen MI, Gisbertz SS. Distribution of lymph node metastases in esophageal carcinoma patients undergoing upfront surgery: a systematic review. *Cancers (Basel)*. 2020;12(6):1592. doi:[10.3390/cancers12061592](https://doi.org/10.3390/cancers12061592)
24. Cuesta MA, van der Peet DL, Gisbertz SS, Straatman J. Mediastinal lymphadenectomy for esophageal cancer: differences between two countries, Japan and the Netherlands. *Ann Gastroenterol Surg*. 2018;2(3):176-181. doi:[10.1002/ags3.12172](https://doi.org/10.1002/ags3.12172)
25. Eddahchouri Y, van Workum F, van den Wildenberg FJ, et al. European consensus on essential steps of minimally invasive Ivor Lewis and McKeown esophagectomy through Delphi methodology. *Surg Endosc*. 2022;36(1):446-460. doi:[10.1007/s00464-021-08304-5](https://doi.org/10.1007/s00464-021-08304-5)
26. Ubels S, Verstegen M, Klarenbeek B, et al; TENTACLE—Esophagus Collaborative Group. Severity of oesophageal anastomotic leak in patients after oesophagectomy: the SEAL score. *Br J Surg*. 2022;109(9):864-871. doi:[10.1093/bjs/znac226](https://doi.org/10.1093/bjs/znac226)
27. Busweiler LA, Schouwenburg MG, van Berge Henegouwen MI, et al; Dutch Upper Gastrointestinal Cancer Audit (DUCA) Group. Textbook outcome as a composite measure in oesophagogastric cancer surgery. *Br J Surg*. 2017;104(6):742-750. doi:[10.1002/bjs.10486](https://doi.org/10.1002/bjs.10486)
28. Zhang J, Yu KF. What's the relative risk? a method of correcting the odds ratio in cohort studies of common outcomes. *JAMA*. 1998;280(19):1690-1691. doi:[10.1001/jama.280.19.1690](https://doi.org/10.1001/jama.280.19.1690)
29. van Workum F, Fransen L, Luyer MD, Rosman C. Learning curves in minimally invasive esophagectomy. *World J Gastroenterol*. 2018;24(44):4974-4978. doi:[10.3748/wjg.v24.i44.4974](https://doi.org/10.3748/wjg.v24.i44.4974)
30. Haverkamp L, Seesing MFJ, Ruurda JP, Boone J, Hillegersberg RV. Worldwide trends in surgical techniques in the treatment of esophageal and gastroesophageal junction cancer. *Dis Esophagus*. 2017;30(1):1-7.

31. Bonrath EM, Dedy NJ, Gordon LE, Grantcharov TP. Comprehensive surgical coaching enhances surgical skill in the operating room: a randomized controlled trial. *Ann Surg*. 2015;262(2):205-212. doi:10.1097/SLA.0000000000001214
32. Soucisse ML, Boulva K, Sideris L, Drolet P, Morin M, Dubé P. Video coaching as an efficient teaching method for surgical residents—a randomized controlled trial. *J Surg Educ*. 2017;74(2):365-371. doi:10.1016/j.jsurg.2016.09.002
33. McQueen S, McKinnon V, VanderBeek L, McCarthy C, Sonnadara R. Video-based assessment in surgical education: a scoping review. *J Surg Educ*. 2019;76(6):1645-1654. doi:10.1016/j.jsurg.2019.05.013
34. Daniel R, McKechnie T, Kruse CC, et al. Video-based coaching for surgical residents: a systematic review and meta-analysis. *Surg Endosc*. 2023;37(2):1429-1439. doi:10.1007/s00464-022-09379-4
35. Igaki T, Kitaguchi D, Matsuzaki H, et al. Automatic surgical skill assessment system based on concordance of standardized surgical field development using artificial intelligence. *JAMA Surg*. 2023;158(8):e231131. doi:10.1001/jamasurg.2023.1131
36. Wu J, Hines OJ. Using artificial intelligence to assess surgeon skill. *JAMA Surg*. 2023;158(8):e231140. doi:10.1001/jamasurg.2023.1140

SUPPLEMENT 1.

eMethods 1. MIE-CAT Scores Calculation

eMethods 2. Multilevel Confounders

eTable 1. Characteristics of the Participating Hospitals

eTable 2. Overview of Surgeons in Participating Centers

eTable 3. Patient Characteristics

eTable 4. Multilevel Analysis of Oncologic Results

eTable 5. Additional Multilevel Analysis Results

eTable 6. Performance Scores Between International and Dutch Experts

eFigure 1. Performance vs Experience (No. of MIEs Over 2 Years)

eFigure 2. Performance vs Volume (No. of MIE Cases Over 2 Years)

eFigure 3. Total MIE-CAT Score Per Operation Type

eAppendix. MIE-CAT Netherlands Collaborative Group

SUPPLEMENT 2.

Nonauthor Collaborators

SUPPLEMENT 3.

Data Sharing Statement