

Supplementary Material

A. Postoperative Complications

In our study, we define a “postoperative complication” in a patient as the occurrence of at least one of the following adverse events after lung cancer surgery:

- Adult Respiratory Distress Syndrome
- Pneumonia
- Atelectasis Requiring Bronchoscopy
- Bronchopleural Fistula
- Pneumothorax
- Air Leak Greater Than Five Days
- Tracheostomy
- Unexpected Admission To ICU
- Empyema Requiring Treatment
- Initial Vent Support > 48 Hours

B. Preoperative Features

B.1. Clinical Features

MIRACLE takes 17 preoperative clinical features of a patient as input, divided into 8 categories, defined as follows:

- General EHR data: Age, BMI, Gender
- Preoperative Pulmonary Functional Tests: Forced Expiratory Volume in 1 second (FEV1) and Diffusing capacity of the lungs for carbon monoxide (DLCO)
- Prior Surgeries: Prior Cardiothoracic Surgery, Preoperative Chemo - Current Malignancy and Preoperative Thoracic Radiation Therapy
- Smoking Status quantifiers: Cigarette Smoking status and Pack-Years Of Cigarette Use
- Fitness indicators: ECOG score and ASA classification
- Tumor Size: Diameter of tumor in centimeters
- TNM classification: Tumor Node Metastasis stage
- Procedure type to be performed: Type of procedure that will take during surgery which can be Chest Wall Excision, Decortification, Lobectomy, wedge resection, etc.

B.2. Radiological Features

MIRACLE takes CT scans of patients as input, which are then processed through Total Segmentator and Pyradiomics, extracting around 113 features, divided into 8 categories, defined as follows:

- First order statistics
- Shape-based features (3D)

- Shape-based features (2D)
- Gray Level Co-occurrence Matrix
- Gray Level Run Length Matrix
- Gray Level Size Zone Matrix
- Neighbouring Gray Tone Difference Matrix
- Gray Level Dependence Matrix

C. Additional examples to examine the quality of LLM-Generated Remarks

In this section, we present more qualitative assessment of the textual remarks produced by each LLM variant as compared to those written by expert surgeons. Our goal is to illustrate typical strengths and failure modes of the generated remarks from LLM with the use of more examples. From the 200-patient test set, we sampled few more cases per category for each of the three LLMs (Llama 3.3 70B-Instruct, DeepSeek R1-Distill Qwen-32B, OpenBioLLM-70B). An expert lung cancer surgeon (not included in the pool of surgeon who wrote the surgeon remarks for the test set) served as an impartial adjudicator to compare, on each case, the surgeon-written remark against the model-generated remark. We categorized the outcome into three categories as done in the main text.

C.1. Examples

C.1.1. LLM Performs Better Than Surgeon

Below are few representative cases per LLM where the model’s remark added new insight or corrected an omission in the surgeon’s comment. We provide the clinical summary of the patient (which acts as input) followed by Surgeon’s remark and remark from LLM.

DeepSeek R1-Distill Qwen-32B

- **Case 1: Clinical Summary:** “Patient is a 66.0-year-old female with a BMI of 22.10176203. Her pulmonary function tests show a predicted FEV1 of 78.0% and a predicted DLCO of 73.0%.She has a smoking history of 22 pack-years, and is classified as a past smoker.Patient has no history of prior cardiothoracic surgery, never undergone preoperative chemotherapy for a current malignancy and has not received preoperative thoracic radiation therapy.Her ECOG performance score is 0, indicating she is

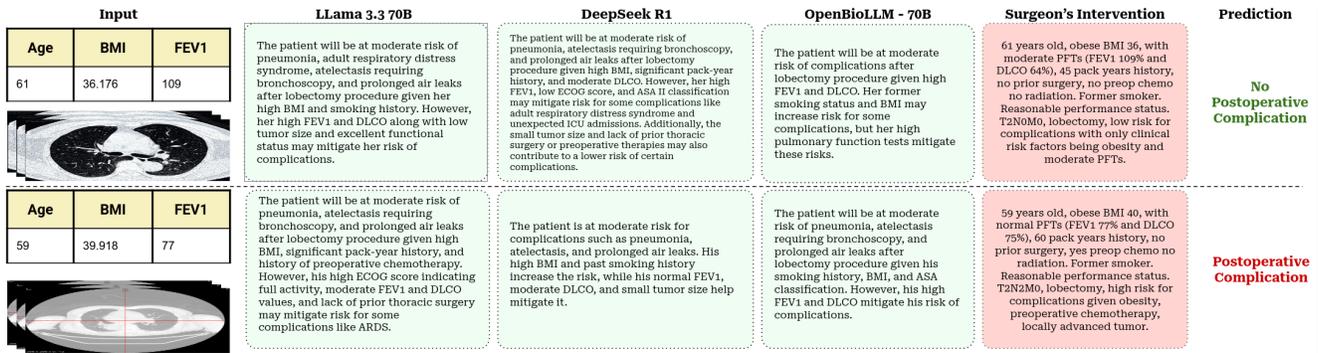


Figure 1. Examples from POC-L test set along with its associated surgeon's remark and remarks generated by all 3 LLMs. (Upper) An Example in which all LLM remarks were better than that of surgeon and (Lower) an example in which surgeon's remark is better than that generated by all LLMs. (Best viewed digitally). More examples are provided in section C.

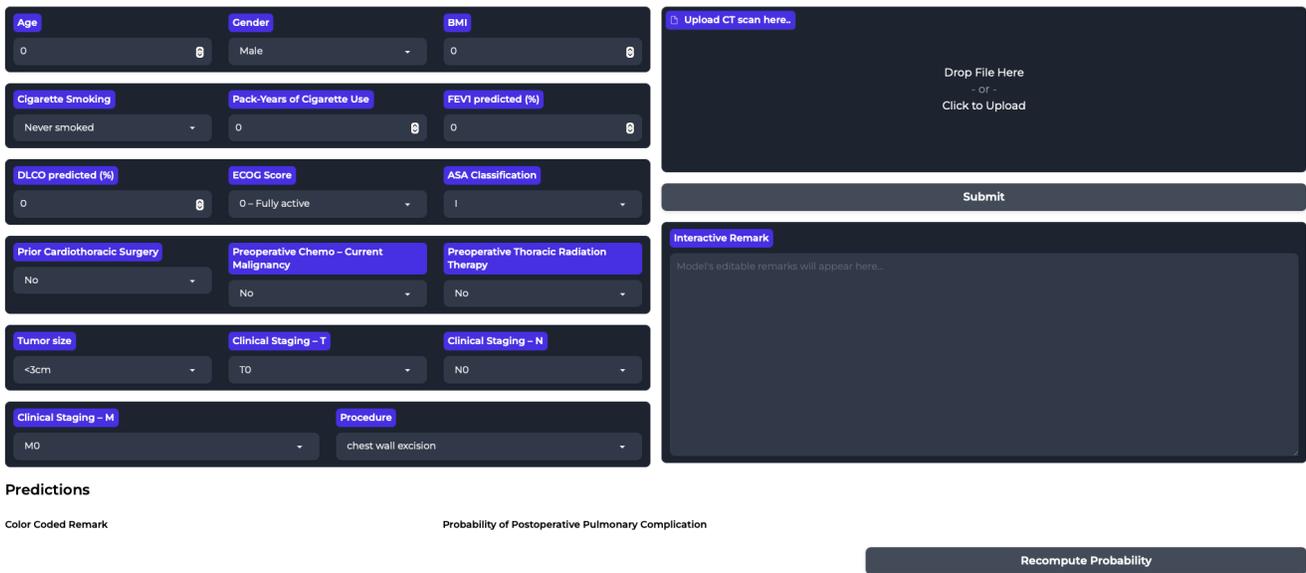


Figure 2. Screenshot of Web based interface built for easy use and external multi-institutional validation of MIRACLE (Best Viewed Digitally)

fully active and able to carry on all pre-disease performance without restriction. She is classified as ASA class II. The tumor is less than 3 centimeter and staged as T1, N0, M0. The planned surgical procedure is a lobectomy.”

Surgeon: “66 year old patient, normal BMI 22, with good PFTs (FEV1 78% and DLCO 73%) and performance status, no prior surgery, no preop chemo or radiation. Former smoker, 22 pack year history. T1N0M0, small tumor and lobectomy, should be low risk for complication.”

LLM: “The patient will be at low risk of postoperative complications such as pneumonia, adult respiratory distress syndrome, atelectasis requiring bronchoscopy, and prolonged air leaks after lobectomy procedure based on her optimal BMI, fully active ECOG score of 0, and ASA

class II. Her moderate FEV1 and DLCO, along with her history of smoking, may slightly increase the risk of some complications, but these risks are mitigated by her overall good functional status and lack of prior thoracic surgery or preoperative therapies.”

- **Case 2: Clinical Summary:** “Patient is a 61.0-year-old female with a BMI of 36.17638311. Her pulmonary function tests show a predicted FEV1 of 109.0% and a predicted DLCO of 64.0%. She has a smoking history of 45 pack-years, and is classified as a past smoker. Patient has no history of prior cardiothoracic surgery, never undergone preoperative chemotherapy for a current malignancy and has not received preoperative thoracic radiation therapy. Her ECOG performance score is 0, indicating she is

fully active and able to carry on all pre-disease performance without restriction. She is classified as ASA class II. The tumor is less than 3 centimeter and staged as T2, N0, M0. The planned surgical procedure is a lobectomy.”

Surgeon: “61 years old, obese BMI 36, with moderate PFTs (FEV1 109% and DLCO 64%), 45 pack years history, no prior surgery, no preop chemo no radiation. Former smoker. Reasonable performance status. T2N0M0, lobectomy, low risk for complications with only clinical risk factors being obesity and moderate PFTs.”

LLM: “The patient will be at moderate risk of pneumonia, adult respiratory distress syndrome, atelectasis requiring bronchoscopy, and prolonged air leaks after lobectomy procedure given her high BMI and smoking history. However, her high FEV1 and DLCO along with low tumor size and excellent functional status may mitigate her risk of complications.”

- **Case 3: Clinical Summary:** “Patient is a 56.0-year-old female with a BMI of 33.8813735691988. Her pulmonary function tests show a predicted FEV1 of 52.0% and a predicted DLCO of 79.0%. She has a smoking history of 30 pack-years and is currently a smoker. Patient has no history of prior cardiothoracic surgery, never undergone preoperative chemotherapy for a current malignancy and has not received preoperative thoracic radiation therapy. Her ECOG performance score is 0, indicating she is fully active and able to carry on all pre-disease performance without restriction. She is classified as ASA class III. The tumor is less than 3 centimeter and staged as T1, N0, M0. The planned surgical procedure is a wedge.”

Surgeon: “56 years old, obese BMI 34, with moderately severe PFTs (FEV1 52% and DLCO 79%), 30 pack years history, no prior surgery, no preop chemo no radiation. Current smoker. Reasonable performance status. T1N0M0, wedge, low risk for complications with only risk factors obesity and moderately severe PFTs, but sublobar resection which is lower risk.”

LLM: “The patient is at moderate risk of postoperative complications such as pneumonia, atelectasis requiring bronchoscopy, and prolonged air leaks. This is primarily due to her high BMI and current smoking status, which increase the likelihood of respiratory complications. However, her good DLCO, absence of prior thoracic surgery, and excellent functional status (ECOG 0) mitigate some of these risks. The less invasive nature of the wedge resection further reduces the risk compared to more extensive procedures.”

- **Case 4: Clinical Summary:** “Patient is a 75.0-year-old female with a BMI of 27.3120605732829. Her pulmonary function tests show a predicted FEV1 of 56.0% and a predicted DLCO of 41.0%. She has a smoking history of 55 pack-years, and is classified as a past smoker. Patient has a history of prior cardiothoracic

surgery, received preoperative thoracic radiation therapy but not undergoing preoperative chemotherapy for a current malignancy. Her ECOG performance score is 0, indicating she is fully active and able to carry on all pre-disease performance without restriction. She is classified as ASA class III. The tumor is less than 3 centimeter and staged as T1, N0, M0. The planned surgical procedure is a wedge.”

Surgeon: “75 years old, overweight BMI 27, with severe PFTs (FEV1 56% and DLCO 41%), 55 pack years history, yes prior surgery, no preop chemo yes radiation. Former smoker. Reasonable performance status. T1N0M0, wedge, moderate to high risk for complications given severe PFTs, prior surgery and radiation.”

LLM: “The patient is at moderate risk for postoperative complications. Her age and prior surgery increase the risk, particularly for pneumonia and atelectasis. However, her optimal BMI, fully active status (ECOG 0), and the minimally invasive nature of the wedge resection mitigate these risks. The DLCO is borderline low, which could slightly elevate the risk of ARDS, but her moderate FEV1 and lack of current smoking help reduce this risk. Overall, the combination of her age and prior surgery contribute to a moderate risk profile, while her good functional status and procedure type lower the likelihood of severe complications.”

Llama 3.3 70B-Instruct

- **Case 1: Clinical Summary:** “Patient is a 47.0-year-old male with a BMI of 19.87654321. His pulmonary function tests show a predicted FEV1 of 90.0% and a predicted DLCO of 89.0%. He has a smoking history of 18 pack-years and is currently a smoker. Patient has no history of prior cardiothoracic surgery, never undergone preoperative chemotherapy for a current malignancy but received preoperative thoracic radiation therapy. His ECOG performance score is 1, indicating he is restricted in physically strenuous activity but is ambulatory and able to carry out work of a light or sedentary nature. He is classified as ASA class III. The tumor is greater than 7 centimeter and staged as T4, N1, M0. The planned surgical procedure is a lobectomy.”

Surgeon: “47 years old, normal BMI 20, with good PFTs (FEV1 90% and DLCO 89%), 18 pack year history, no prior surgery, no preop chemo, did have preop radiation. Current smoker. Good performance status. Large tumor (over 7cm) T4N1M0, lobectomy, should be moderate risk for complication.”

LLM: “The patient will be at high risk of pneumonia, adult respiratory distress syndrome, atelectasis requiring bronchoscopy, and prolonged air leaks after lobectomy procedure given low BMI, current smoking status, history of preoperative thoracic radiation therapy, and high

tumor stage. However, his high FEV1 and DLCO may mitigate his risk for some complications like unexpected ICU admissions and initial vent support greater than 48 hours. Additionally, his ECOG score of 1 and relatively low pack-year history may also contribute to a slightly lower risk for certain complications, but the overall high-risk profile due to low BMI, smoking, and advanced tumor stage predominates.”

- **Case 2: Clinical Summary:** “Patient is a 67.0-year-old female with a BMI of 31.05098855359. Her pulmonary function tests show a predicted FEV1 of 75.0% and a predicted DLCO of 50.0%. She has a smoking history of 35 pack-years and is currently a smoker. Patient has no history of prior cardiothoracic surgery, never undergone preoperative chemotherapy for a current malignancy and has not received preoperative thoracic radiation therapy. Her ECOG performance score is 0, indicating she is fully active and able to carry on all pre-disease performance without restriction. She is classified as ASA class III. The tumor is between 3 to 5 centimeter and staged as T2, N0, M0. The planned surgical procedure is a lobectomy.”

Surgeon: “67 years old, overweight BMI 31, with moderate PFTs (FEV1 75% and DLCO 50%), 35 pack year history, no prior surgery, no preop chemo or radiation. Current smoker. Reasonable performance status. T2N0M0, lobectomy, should be moderate risk for complication given moderate PFTs and current smoker.”

LLM: “The patient will be at moderate risk of pneumonia, atelectasis requiring bronchoscopy, and prolonged air leaks after lobectomy procedure given current smoking status, high pack-year history, and moderate DLCO. Her high BMI may increase risk for infections and delayed recovery, while her high FEV1 and good ECOG performance score may mitigate risk for some complications like adult respiratory distress syndrome and unexpected ICU admissions. Additionally, the planned lobectomy procedure and tumor size between 3 to 5 centimeters may also contribute to the moderate risk of complications.”

- **Case 3: Clinical Summary:** “Patient is a 64.0-year-old female with a BMI of 28.32166449. Her pulmonary function tests show a predicted FEV1 of 95.0% and a predicted DLCO of 52.0%. She has a smoking history of 45 pack-years, and is classified as a past smoker. Patient has no history of prior cardiothoracic surgery, never undergone preoperative chemotherapy for a current malignancy and has not received preoperative thoracic radiation therapy. Her ECOG performance score is 0, indicating she is fully active and able to carry on all pre-disease performance without restriction. She is classified as ASA class III. The tumor is between 3 to 5 centimeter and staged as T2, N0, M0. The planned surgical procedure is a lobectomy.”

Surgeon: “64 years old, normal BMI 28, with moder-

ately severe PFTs (FEV1 95% and DLCO 52%), 45 pack year history, no prior surgery, no preop chemo no radiation. Former smoker. Reasonable performance status. T2N0M0, lobectomy, moderate risk for complications given PFTs.”

LLM: “The patient will be at moderate risk of pneumonia, atelectasis requiring bronchoscopy, and prolonged air leaks after lobectomy procedure given her moderate to high pack-year history, ASA III classification, and relatively low DLCO. However, her high FEV1, ECOG score of 0, and lack of prior thoracic surgery may mitigate risk for some complications like adult respiratory distress syndrome and unexpected ICU admissions.”

- **Case 4: Clinical Summary:** “Patient is a 61.0-year-old female with a BMI of 36.17638311. Her pulmonary function tests show a predicted FEV1 of 109.0% and a predicted DLCO of 64.0%. She has a smoking history of 45 pack-years, and is classified as a past smoker. Patient has no history of prior cardiothoracic surgery, never undergone preoperative chemotherapy for a current malignancy and has not received preoperative thoracic radiation therapy. Her ECOG performance score is 0, indicating she is fully active and able to carry on all pre-disease performance without restriction. She is classified as ASA class II. The tumor is less than 3 centimeter and staged as T2, N0, M0. The planned surgical procedure is a lobectomy.”

Surgeon: “61 years old, obese BMI 36, with moderate PFTs (FEV1 109% and DLCO 64%), 45 pack years history, no prior surgery, no preop chemo no radiation. Former smoker. Reasonable performance status. T2N0M0, lobectomy, low risk for complications with only clinical risk factors being obesity and moderate PFTs.”

LLM: “The patient will be at moderate risk of pneumonia, atelectasis requiring bronchoscopy, and prolonged air leaks after lobectomy procedure given high BMI, significant pack-year history, and moderate DLCO. However, her high FEV1, low ECOG score, and ASA II classification may mitigate risk for some complications like adult respiratory distress syndrome and unexpected ICU admissions. Additionally, the small tumor size and lack of prior thoracic surgery or preoperative therapies may also contribute to a lower risk of certain complications.”

OpenBioLLM-70B

- **Case 1: Clinical Summary:** “Patient is a 64.0-year-old female with a BMI of 36.06637186. Her pulmonary function tests show a predicted FEV1 of 85.0% and a predicted DLCO of 60.0%. She has a smoking history of 49 pack-years and is currently a smoker. Patient has no history of prior cardiothoracic surgery, never undergone preoperative chemotherapy for a current malignancy and has not received preoperative thoracic radiation therapy. Her ECOG performance score is 1, indicating she is restricted

in physically strenuous activity but is ambulatory and able to carry out work of a light or sedentary nature. She is classified as ASA class III. The tumor is less than 3 centimeter and staged as T1, N0, M0. The planned surgical procedure is a wedge.”

Surgeon: “64 years old, obese BMI 36, with moderate PFTs (FEV1 85% and DLCO 60%), 49 pack years history, no prior surgery, no preop chemo no radiation. Current smoker. Reasonable performance status. T1N0M0, wedge, low to moderate risk for complications with current smoker and obese, but sublobar resection should lower the risk.”

LLM: “The patient will be at moderate risk of pneumonia, atelectasis requiring bronchoscopy, and prolonged air leaks after wedge procedure given her current smoking status, low BMI, and history of pack-years. However, her high FEV1 and DLCO mitigate her risk of complications.”

- **Case 2: Clinical Summary:** “Patient is a 56.0-year-old male with a BMI of 27.95493197. His pulmonary function tests show a predicted FEV1 of 63.0% and a predicted DLCO of 82.0%. He has a smoking history of 35 pack-years, and is classified as a past smoker. Patient has no history of prior cardiothoracic surgery, never undergone preoperative chemotherapy for a current malignancy and has not received preoperative thoracic radiation therapy. His ECOG performance score is 0, indicating he is fully active and able to carry on all pre-disease performance without restriction. He is classified as ASA class II. The tumor is less than 3 centimeter and staged as T1, N0, M0. The planned surgical procedure is a lobectomy.”

Surgeon: “56 years old, overweight BMI 28, with moderate PFTs (FEV1 63% and DLCO 82%), 35 pack years history, no prior surgery, no preop chemo no radiation. Former smoker. Reasonable performance status. T1N0M0, lobectomy, low risk for complications given clinical data as PFTs moderate, former smoker, no neoadjuvant treatment, small tumor.”

LLM: “The patient will be at moderate risk of pneumonia, atelectasis requiring bronchoscopy, and prolonged air leaks after lobectomy procedure given moderate FEV1 and DLCO, former smoking status, and low pack years. However, his age, BMI, ECOG score, and ASA classification mitigate his risk of complication.”

- **Case 3: Clinical Summary:** “Patient is a 74.0-year-old female with a BMI of 26.65422532. Her pulmonary function tests show a predicted FEV1 of 87.0% and a predicted DLCO of 77.0%. She has a smoking history of 62 pack-years and is currently a smoker. Patient has no history of prior cardiothoracic surgery, never undergone preoperative chemotherapy for a current malignancy and has not received preoperative thoracic radiation therapy. Her ECOG performance score is 0, indicating she is fully

active and able to carry on all pre-disease performance without restriction. She is classified as ASA class III. The tumor is less than 3 centimeter and staged as T1, N0, M0. The planned surgical procedure is a lobectomy.”

Surgeon: “74 years old, overweight BMI 27, with normal PFTs (FEV1 87% and DLCO 77%), 62 pack years history, no prior surgery, no preop chemo no radiation. Current smoker. Reasonable performance status. T1N1M0, lobectomy, low to moderate risk for complications given current smoker.”

LLM: “The patient will be at moderate risk of pneumonia, adult respiratory distress syndrome, atelectasis requiring bronchoscopy, and prolonged air leaks after lobectomy procedure given her current smoking status, pack-year history, and ASA class III. However, her high FEV1 and DLCO mitigate her risk of complications.”

- **Case 4: Clinical Summary:** “Patient is a 52.0-year-old female with a BMI of 21.7079889807162. Her pulmonary function tests show a predicted FEV1 of 94.0% and a predicted DLCO of 90.0%. She has no pack-years of cigarette use, indicating she has never smoked. Patient has a history of prior cardiothoracic surgery, undergoing preoperative chemotherapy for a current malignancy but did not receive preoperative thoracic radiation therapy. Her ECOG performance score is 0, indicating she is fully active and able to carry on all pre-disease performance without restriction. She is classified as ASA class II. The tumor is less than 3 centimeter and staged as T0, N0, M0. The planned surgical procedure is a wedge.”

Surgeon: “52 years old, normal BMI 22, with normal PFTs (FEV1 94% and DLCO 90%), 0 pack years history, yes prior surgery, no preop chemo no radiation. Never smoker. Reasonable performance status. T0N0M0, wedge, low to moderate risk given prior surgery, but young, never smoker, no preoperative treatment, and just a wedge.”

LLM: “The patient will be at low risk of postoperative complications after wedge procedure given high FEV1, high DLCO, and never smoking status. Prior cardiothoracic surgery and ASA class II may increase risk for some complications, but these risks are mitigated by her optimal pulmonary function.”

C.1.2. LLM Comparable to Surgeon

In these few cases, the strength and level of detail in the model’s explanation match that of the surgeon’s, with no clear advantage either way.

DeepSeek R1-Distill Qwen-32B

- **Case 1: Clinical Summary:** “Patient is a 69.0-year-old female with a BMI of 23.62754767. Her pulmonary function tests show a predicted FEV1 of 107.0% and a predicted DLCO of 71.0%. She has a smoking history of 15

pack-years, and is classified as a past smoker. Patient has no history of prior cardiothoracic surgery, never undergone preoperative chemotherapy for a current malignancy and has not received preoperative thoracic radiation therapy. Her ECOG performance score is 0, indicating she is fully active and able to carry on all pre-disease performance without restriction. She is classified as ASA class II. The tumor is less than 3 centimeter and staged as T1, N0, M0. The planned surgical procedure is a lobectomy.”

Surgeon: “69 years old, normal BMI 24, with good PFTs (FEV1 107% and DLCO 71%), 15 pack year history, no prior surgery, no preop chemo no radiation. Former smoker. Reasonable performance status. T1N0M0, lobectomy, low risk for complications.”

LLM: “The patient will be at low risk of postoperative complications after lobectomy procedure given her optimal BMI, excellent pulmonary function (FEV1 107%, DLCO 71%), and excellent functional status (ECOG 0). Her past smoking history and tumor size (>3 cm) may slightly increase the risk of minor complications like atelectasis or air leaks, but these risks are mitigated by her overall good health and pulmonary reserve.”

- **Case 2: Clinical Summary:** “Patient is a 67.0-year-old female with a BMI of 23.95209581. Her pulmonary function tests show a predicted FEV1 of 100.0% and a predicted DLCO of 87.0%. She has a smoking history of 50 pack-years and is currently a smoker. Patient has no history of prior cardiothoracic surgery, never undergone preoperative chemotherapy for a current malignancy and has not received preoperative thoracic radiation therapy. Her ECOG performance score is 0, indicating she is fully active and able to carry on all pre-disease performance without restriction. She is classified as ASA class II. The tumor is less than 3 centimeter and staged as T1, N0, M0. The planned surgical procedure is a lobectomy.”

Surgeon: “67 years old, normal BMI 24, with good PFTs (FEV1 100% and DLCO 87%), 50 pack year history, no prior surgery, no preop chemo no radiation. Current smoker. Reasonable performance status. T1N0M0, lobectomy, moderate risk for complications given current smoker.”

LLM: “The patient is at moderate risk for postoperative complications such as pneumonia and prolonged air leaks due to her smoking history. However, her excellent pulmonary function and overall health mitigate these risks, placing her at a lower risk for severe complications like ARDS or prolonged ICU stay.”

- **Case 3: Clinical Summary:** “Patient is a 72.0-year-old female with a BMI of 23.7386621315193. Her pulmonary function tests show a predicted FEV1 of 111.0% and a predicted DLCO of 78.0%. She has a smoking history of 51 pack-years, and is classified as a past smoker. Patient has no history of prior cardiothoracic

surgery, never undergone preoperative chemotherapy for a current malignancy and has not received preoperative thoracic radiation therapy. Her ECOG performance score is 1, indicating she is restricted in physically strenuous activity but is ambulatory and able to carry out work of a light or sedentary nature. She is classified as ASA class III. The tumor is less than 3 centimeter and staged as T1, N0, M0. The planned surgical procedure is a lobectomy.”

Surgeon: “72 years old, normal BMI 24, with normal PFTs (FEV1 111% and DLCO 78%), 51 pack year history, no prior surgery, no preop chemo no radiation. Former smoker. Reasonable performance status. T1N0M0, lobectomy, moderate risk for complications given poor PFTs but only a wedge procedure.”

LLM: “The patient is at moderate risk for postoperative complications such as pneumonia, atelectasis, and prolonged air leaks. Her advanced age and smoking history contribute to this risk, while her optimal BMI, good pulmonary function, and lack of prior surgery mitigate some risks. The ASA III and ECOG 1 scores further support a moderate risk profile.”

Llama 3.3 70B-Instruct

- **Case 1: Clinical Summary:** “Patient is a 69.0-year-old female with a BMI of 23.62754767. Her pulmonary function tests show a predicted FEV1 of 107.0% and a predicted DLCO of 71.0%. She has a smoking history of 15 pack-years, and is classified as a past smoker. Patient has no history of prior cardiothoracic surgery, never undergone preoperative chemotherapy for a current malignancy and has not received preoperative thoracic radiation therapy. Her ECOG performance score is 0, indicating she is fully active and able to carry on all pre-disease performance without restriction. She is classified as ASA class II. The tumor is less than 3 centimeter and staged as T1, N0, M0. The planned surgical procedure is a lobectomy.”

Surgeon: “69 years old, normal BMI 24, with good PFTs (FEV1 107% and DLCO 71%), 15 pack year history, no prior surgery, no preop chemo no radiation. Former smoker. Reasonable performance status. T1N0M0, lobectomy, low risk for complications.”

LLM: “The patient will be at low risk of pneumonia, adult respiratory distress syndrome, atelectasis requiring bronchoscopy, and prolonged air leaks after lobectomy procedure given high FEV1, moderate DLCO, low BMI within normal range, and low tumor size. Her former smoking status with 15 pack-year history and moderate ASA classification may slightly increase risk for some complications like pneumonia and atelectasis, but high FEV1 and ECOG score of 0 may mitigate these risks.”

- **Case 2: Clinical Summary:** “Patient is a 67.0-year-old female with a BMI of 23.95209581. Her pulmonary function tests show a predicted FEV1 of 100.0% and a pre-

dicted DLCO of 87.0%.She has a smoking history of 50 pack-years and is currently a smoker.Patient has no history of prior cardiothoracic surgery, never undergone preoperative chemotherapy for a current malignancy and has not received preoperative thoracic radiation therapy.Her ECOG performance score is 0, indicating she is fully active and able to carry on all pre-disease performance without restriction.She is classified as ASA class II.The tumor is less than 3 centimeter and staged as T1, N0, M0.The planned surgical procedure is a lobectomy.”

Surgeon: “67 years old, normal BMI 24, with good PFTs (FEV1 100% and DLCO 87%), 50 pack year history, no prior surgery, no preop chemo no radiation. Current smoker. Reasonable performance status. T1N0M0, lobectomy, moderate risk for complications given current smoker.”

LLM: “The patient will be at moderate risk of pneumonia, atelectasis requiring bronchoscopy, and prolonged air leaks after lobectomy procedure given current smoking status and high pack-year history. However, her high FEV1 and DLCO, low BMI within normal range, and good ECOG performance score may mitigate risk for some complications like adult respiratory distress syndrome and unexpected ICU admissions.”

- **Case 3: Clinical Summary:** “Patient is a 72.0-year-old female with a BMI of 23.7386621315193. Her pulmonary function tests show a predicted FEV1 of 111.0% and a predicted DLCO of 78.0%.She has a smoking history of 51 pack-years, and is classified as a past smoker.Patient has no history of prior cardiothoracic surgery, never undergone preoperative chemotherapy for a current malignancy and has not received preoperative thoracic radiation therapy.Her ECOG performance score is 1, indicating she is restricted in physically strenuous activity but is ambulatory and able to carry out work of a light or sedentary nature.She is classified as ASA class III.The tumor is less than 3 centimeter and staged as T1, N0, M0.The planned surgical procedure is a lobectomy.”

Surgeon: “72 years old, normal BMI 24, with normal PFTs (FEV1 111% and DLCO 78%), 51 pack year history, no prior surgery, no preop chemo no radiation. Former smoker. Reasonable performance status. T1N0M0, lobectomy, moderate risk for complications given poor PFTs but only a wedge procedure.”

LLM: “The patient will be at moderate risk of pneumonia, atelectasis requiring bronchoscopy, and unexpected ICU admissions after lobectomy procedure given advanced age, high pack-year history, and ASA III classification. However, her high FEV1 and relatively preserved DLCO may mitigate risk for some complications like adult respiratory distress syndrome and prolonged ventilator support, and her low BMI within the normal range, ECOG score of 1, and small tumor size may also contribute to a

lower risk of certain complications.”

OpenBioLLM-70B

- **Case 1: Clinical Summary:** “Patient is a 69.0-year-old female with a BMI of 23.62754767. Her pulmonary function tests show a predicted FEV1 of 107.0% and a predicted DLCO of 71.0%.She has a smoking history of 15 pack-years, and is classified as a past smoker.Patient has no history of prior cardiothoracic surgery, never undergone preoperative chemotherapy for a current malignancy and has not received preoperative thoracic radiation therapy.Her ECOG performance score is 0, indicating she is fully active and able to carry on all pre-disease performance without restriction.She is classified as ASA class II.The tumor is less than 3 centimeter and staged as T1, N0, M0.The planned surgical procedure is a lobectomy.”

Surgeon: “69 years old, normal BMI 24, with good PFTs (FEV1 107% and DLCO 71%), 15 pack year history, no prior surgery, no preop chemo no radiation. Former smoker. Reasonable performance status. T1N0M0, lobectomy, low risk for complications.”

LLM: “The patient will be at low risk of pneumonia, adult respiratory distress syndrome, atelectasis requiring bronchoscopy, and prolonged air leaks after lobectomy procedure given high FEV1 and DLCO, low pack years, and absence of preoperative thoracic radiation therapy. However, her age, BMI, and smoking history may moderately increase her risk for some complications.”

- **Case 2: Clinical Summary:** “Patient is a 67.0-year-old female with a BMI of 23.95209581. Her pulmonary function tests show a predicted FEV1 of 100.0% and a predicted DLCO of 87.0%.She has a smoking history of 50 pack-years and is currently a smoker.Patient has no history of prior cardiothoracic surgery, never undergone preoperative chemotherapy for a current malignancy and has not received preoperative thoracic radiation therapy.Her ECOG performance score is 0, indicating she is fully active and able to carry on all pre-disease performance without restriction.She is classified as ASA class II.The tumor is less than 3 centimeter and staged as T1, N0, M0.The planned surgical procedure is a lobectomy.”

Surgeon: “67 years old, normal BMI 24, with good PFTs (FEV1 100% and DLCO 87%), 50 pack year history, no prior surgery, no preop chemo no radiation. Current smoker. Reasonable performance status. T1N0M0, lobectomy, moderate risk for complications given current smoker.”

LLM: “The patient will be at moderate risk of pneumonia, atelectasis requiring bronchoscopy, and prolonged air leaks after lobectomy procedure given current smoking status and history of pack years. However, her high FEV1 and DLCO mitigate her risk of complication.”

- **Case 3: Clinical Summary:** “Patient is a 61.0-year-old

female with a BMI of 36.17638311. Her pulmonary function tests show a predicted FEV1 of 109.0% and a predicted DLCO of 64.0%. She has a smoking history of 45 pack-years, and is classified as a past smoker. Patient has no history of prior cardiothoracic surgery, never undergone preoperative chemotherapy for a current malignancy and has not received preoperative thoracic radiation therapy. Her ECOG performance score is 0, indicating she is fully active and able to carry on all pre-disease performance without restriction. She is classified as ASA class II. The tumor is less than 3 centimeter and staged as T2, N0, M0. The planned surgical procedure is a lobectomy.”

Surgeon: “61 years old, obese BMI 36, with moderate PFTs (FEV1 109% and DLCO 64%), 45 pack years history, no prior surgery, no preop chemo no radiation. Former smoker. Reasonable performance status. T2N0M0, lobectomy, low risk for complications with only clinical risk factors being obesity and moderate PFTs.”

LLM: “The patient will be at low risk of complications after lobectomy procedure given high FEV1 and DLCO. Her former smoking status and BMI may increase risk for some complications, but her high pulmonary function tests mitigate these risks.”

C.1.3. LLM Performs Worse Than Surgeon

These few examples highlight situations where the model’s remark omitted a key clinical factor or exhibited an incorrect inference.

DeepSeek R1-Distill Qwen-32B

- **Case 1: Clinical Summary:** “Patient is a 70.0-year-old female with a BMI of 21.9532455203513. Her pulmonary function tests show a predicted FEV1 of 100.0% and a predicted DLCO of 90.0%. She has a smoking history of 35 pack-years, and is classified as a past smoker. Patient has no history of prior cardiothoracic surgery, never undergone preoperative chemotherapy for a current malignancy and has not received preoperative thoracic radiation therapy. Her ECOG performance score is 0, indicating she is fully active and able to carry on all pre-disease performance without restriction. She is classified as ASA class III. The tumor is between 3 to 5 centimeter and staged as T2, N0, M0. The planned surgical procedure is a lobectomy.”

Surgeon: “70 years old, low BMI 22, with good PFTs (FEV1 100% and DLCO 90%), 35 pack year history, no prior surgery, no preop chemo or radiation. Former smoker. Reasonable performance status. T2N0M0, lobectomy, should be low risk for complications.”

LLM: “The patient is at moderate risk of postoperative complications such as pneumonia, atelectasis requiring bronchoscopy, and prolonged air leaks. This is due to her age and low BMI, which are risk factors, but her excellent pulmonary function (FEV1 and DLCO) and being

a former smoker mitigate some of these risks. Her ASA class III and tumor size add a slight elevation in risk, but overall, her risk for severe complications like ARDS or prolonged ICU stays is lower due to her good pulmonary function.”

- **Case 2: Clinical Summary:** “Patient is a 64.0-year-old female with a BMI of 28.32166449. Her pulmonary function tests show a predicted FEV1 of 95.0% and a predicted DLCO of 52.0%. She has a smoking history of 45 pack-years, and is classified as a past smoker. Patient has no history of prior cardiothoracic surgery, never undergone preoperative chemotherapy for a current malignancy and has not received preoperative thoracic radiation therapy. Her ECOG performance score is 0, indicating she is fully active and able to carry on all pre-disease performance without restriction. She is classified as ASA class III. The tumor is between 3 to 5 centimeter and staged as T2, N0, M0. The planned surgical procedure is a lobectomy.”

Surgeon: “64 years old, normal BMI 28, with moderately severe PFTs (FEV1 95% and DLCO 52%), 45 pack year history, no prior surgery, no preop chemo no radiation. Former smoker. Reasonable performance status. T2N0M0, lobectomy, moderate risk for complications given PFTs.”

LLM: “The patient will be at moderate risk of postoperative complications such as pneumonia, atelectasis requiring bronchoscopy, prolonged air leaks, and unexpected ICU admission. Her high BMI increases the risk of infections and respiratory complications, while her normal FEV1 and moderate DLCO help mitigate some risks. Her former smoking status and ASA III classification further contribute to the moderate risk profile.”

Llama 3.3 70B-Instruct

- **Case 1: Clinical Summary:** “Patient is a 70.0-year-old female with a BMI of 21.9532455203513. Her pulmonary function tests show a predicted FEV1 of 100.0% and a predicted DLCO of 90.0%. She has a smoking history of 35 pack-years, and is classified as a past smoker. Patient has no history of prior cardiothoracic surgery, never undergone preoperative chemotherapy for a current malignancy and has not received preoperative thoracic radiation therapy. Her ECOG performance score is 0, indicating she is fully active and able to carry on all pre-disease performance without restriction. She is classified as ASA class III. The tumor is between 3 to 5 centimeter and staged as T2, N0, M0. The planned surgical procedure is a lobectomy.”

Surgeon: “70 years old, low BMI 22, with good PFTs (FEV1 100% and DLCO 90%), 35 pack year history, no prior surgery, no preop chemo or radiation. Former smoker. Reasonable performance status. T2N0M0,

lobectomy, should be low risk for complications.”

LLM: “The patient will be at moderate risk of pneumonia, atelectasis requiring bronchoscopy, and prolonged air leaks after lobectomy procedure given advanced age, history of smoking, and moderate ASA classification. However, her high FEV1 and DLCO, low BMI, and good ECOG performance score may mitigate risk for some complications like adult respiratory distress syndrome and unexpected ICU admissions.”

- **Case 2: Clinical Summary:** “Patient is a 82.0-year-old female with a BMI of 16.75365116. Her pulmonary function tests show a predicted FEV1 of 63.0% and a predicted DLCO of 63.0%.She has a smoking history of 22 pack-years and is currently a smoker.Patient has no history of prior cardiothoracic surgery, never undergone preoperative chemotherapy for a current malignancy and has not received preoperative thoracic radiation therapy.Her ECOG performance score is 0, indicating she is fully active and able to carry on all pre-disease performance without restriction.She is classified as ASA class III.The tumor is less than 3 centimeter and staged as T1, N0, M0.The planned surgical procedure is a lobectomy.”

Surgeon: “82 years old, underweight BMI 17, with moderate PFTs (FEV1 63% and DLCO 63%), 22 pack years history, no prior surgery, no preop chemo no radiation. Current smoker. Reasonable performance status. T1N0M0, lobectomy, moderate to high risk for complications given age, current smoker, depressed PFTs, low weight/frailty, and lobectomy.”

LLM: “The patient will be at high risk of pneumonia, adult respiratory distress syndrome, atelectasis requiring bronchoscopy, and prolonged air leaks after lobectomy procedure given low BMI, current smoking status, and advanced age. Her low BMI and current smoking status may increase risk for some complications like pneumonia and atelectasis requiring bronchoscopy, while her moderate FEV1 and DLCO may mitigate risk for some complications like prolonged ventilator support. However, her high pack-year history and ASA III classification may further increase her risk for severe complications.”

OpenBioLLM-70B

- **Case 1: Clinical Summary:** “Patient is a 82.0-year-old female with a BMI of 16.75365116. Her pulmonary function tests show a predicted FEV1 of 63.0% and a predicted DLCO of 63.0%.She has a smoking history of 22 pack-years and is currently a smoker.Patient has no history of prior cardiothoracic surgery, never undergone preoperative chemotherapy for a current malignancy and has not received preoperative thoracic radiation therapy.Her ECOG performance score is 0, indicating she is fully active and able to carry on all pre-disease performance without restriction.She is classified as ASA class III.The

tumor is less than 3 centimeter and staged as T1, N0, M0.The planned surgical procedure is a lobectomy.”

Surgeon: “82 years old, underweight BMI 17, with moderate PFTs (FEV1 63% and DLCO 63%), 22 pack years history, no prior surgery, no preop chemo no radiation. Current smoker. Reasonable performance status. T1N0M0, lobectomy, moderate to high risk for complications given age, current smoker, depressed PFTs, low weight/frailty, and lobectomy.”

LLM: “The patient will be at high risk of pneumonia, adult respiratory distress syndrome, atelectasis requiring bronchoscopy, and prolonged air leaks after lobectomy procedure given low BMI, current smoking status, and history of preoperative thoracic radiation therapy. However her high FEV1 and DLCO along with low pack years mitigates her risk of complication.”

- **Case 2: Clinical Summary:** “Patient is a 82.0-year-old female with a BMI of 26.23828944. Her pulmonary function tests show a predicted FEV1 of 76.0% and a predicted DLCO of 23.0%.She has a smoking history of 23 pack-years, and is classified as a past smoker.Patient has no history of prior cardiothoracic surgery, never undergone preoperative chemotherapy for a current malignancy and has not received preoperative thoracic radiation therapy.Her ECOG performance score is 0, indicating she is fully active and able to carry on all pre-disease performance without restriction.She is classified as ASA class II.The tumor is less than 3 centimeter and staged as T1, N0, M0.The planned surgical procedure is a lobectomy.”

Surgeon: “82 years old, overweight BMI 26, with very severe PFTs (FEV1 76% and DLCO 23%), 23 pack years history, no prior surgery, no preop chemo no radiation. Former smoker. Reasonable performance status. T1N0M0, lobectomy, High risk for complications given age, severe PFTs, and lobectomy.”

LLM: “The patient will be at moderate risk of pneumonia, atelectasis requiring bronchoscopy, and prolonged air leaks after lobectomy procedure given advanced age, smoking history, and low DLCO. However, her high FEV1 and good ECOG score mitigate her risk of complication.”